

PCTWORLD INTELLECTUAL PROPERTY ORGANIZATION
International Bureau

INTERNATIONAL APPLICATION PUBLISHED UNDER THE PATENT COOPERATION TREATY (PCT)

(51) International Patent Classification ⁶ : A61K 38/00, 45/05, C07K 14/00, 14/82, C12N 15/00, C12Q 1/00		A1	(11) International Publication Number: WO 98/55133
			(43) International Publication Date: 10 December 1998 (10.12.98)
(21) International Application Number: PCT/US98/11533 (22) International Filing Date: 4 June 1998 (04.06.98) (30) Priority Data: 08/870,941 6 June 1997 (06.06.97) US (71) Applicants: THE REGENTS OF THE UNIVERSITY OF CALIFORNIA [US/US]; 12th floor, 1111 Franklin Street, Oakland, CA 94607-5200 (US). UNIVERSITY OF SOUTHERN CALIFORNIA [US/US]; University Park, Los Angeles, CA 90089 (US). (72) Inventors: MITCHELL, Malcolm, S.; P.O. Box 676060, Rancho Santa Fe, CA 92067 (US). DEANS, Robert, J.; 415 Furman Drive, Clairemont, CA 91711 (US). MINEV, Boris, R.; 12496 Cavallo Street, San Diego, CA 92130-2734 (US). KAN-MITCHELL, June; P.O. Box 676060, Rancho Santa Fe, CA 92067 (US). (74) Agents: IMBRA, Richards, J. et al.; Campbell Flores LLP, Suite 700, 4370 La Jolla Village Drive, San Diego, CA 92122 (US).			(81) Designated States: AU, CA, JP, European patent (AT, BE, CH, CY, DE, DK, ES, FI, FR, GB, GR, IE, IT, LU, MC, NL, PT, SE). Published <i>With international search report. Before the expiration of the time limit for amending the claims and to be republished in the event of the receipt of amendments.</i>
(54) Title: A MELANOMA ASSOCIATED ANTIGEN, T CELL EPITOPES THEREOF AND METHODS OF USING SAME			
(57) Abstract The present invention provides a substantially purified polypeptide portion of a melanoma associated antigen, MG50 (SEQ ID NO: 2) and substantially purified T cell epitopes of MG50. For example, the invention provides a cytotoxic T cell epitope having the amino acid sequence RPRPEQUEPLP (SEQ ID NO: 4) and a helper T cell epitope having the amino acid sequence CSEQPFPEHTASVQHAD (SEQ. ID NO: 3). The invention also provides antibodies that specifically bind to MG50 or an MG50 T cell epitope and provides antigen binding fragments of such antibodies. Also provided are a substantially purified nucleic acid molecule (SEQ ID NO: 1), which encodes a portion of a melanoma associated antigen, MG50, and nucleic acid molecules encoding MG50 T cell epitopes. Vectors containing such nucleic acid molecules and cells containing such vectors also are provided. For example, antigen presenting cells containing a nucleic acid molecule of the invention are provided. The invention also provides methods of identifying an MG50 melanoma associated antigen in an individual and methods of identifying an immune response against an MG50 melanoma associated antigen in an individual. In addition, the invention provides methods of stimulating T lymphocytes that are reactive against cancer cells expressing an MG50 melanoma associated antigen and provides methods of treating an individual having cancer cells that express an MG50 melanoma associated antigen.			

FOR THE PURPOSES OF INFORMATION ONLY

Codes used to identify States party to the PCT on the front pages of pamphlets publishing international applications under the PCT.

AL	Albania	ES	Spain	LS	Lesotho	SI	Slovenia
AM	Armenia	FI	Finland	LT	Lithuania	SK	Slovakia
AT	Austria	FR	France	LU	Luxembourg	SN	Senegal
AU	Australia	GA	Gabon	LV	Latvia	SZ	Swaziland
AZ	Azerbaijan	GB	United Kingdom	MC	Monaco	TD	Chad
BA	Bosnia and Herzegovina	GE	Georgia	MD	Republic of Moldova	TG	Togo
BB	Barbados	GH	Ghana	MG	Madagascar	TJ	Tajikistan
BE	Belgium	GN	Guinea	MK	The former Yugoslav Republic of Macedonia	TM	Turkmenistan
BF	Burkina Faso	GR	Greece			TR	Turkey
BG	Bulgaria	HU	Hungary	ML	Mali	TT	Trinidad and Tobago
BJ	Benin	IE	Ireland	MN	Mongolia	UA	Ukraine
BR	Brazil	IL	Israel	MR	Mauritania	UG	Uganda
BY	Belarus	IS	Iceland	MW	Malawi	US	United States of America
CA	Canada	IT	Italy	MX	Mexico	UZ	Uzbekistan
CF	Central African Republic	JP	Japan	NE	Niger	VN	Viet Nam
CG	Congo	KE	Kenya	NL	Netherlands	YU	Yugoslavia
CH	Switzerland	KG	Kyrgyzstan	NO	Norway	ZW	Zimbabwe
CI	Côte d'Ivoire	KP	Democratic People's Republic of Korea	NZ	New Zealand		
CM	Cameroon	KR	Republic of Korea	PL	Poland		
CN	China	KZ	Kazakstan	PT	Portugal		
CU	Cuba	LC	Saint Lucia	RO	Romania		
CZ	Czech Republic	LI	Liechtenstein	RU	Russian Federation		
DE	Germany	LK	Sri Lanka	SD	Sudan		
DK	Denmark	LR	Liberia	SE	Sweden		
EE	Estonia			SG	Singapore		

A MELANOMA ASSOCIATED ANTIGEN, T CELL EPITOPES THEREOF
AND METHODS OF USING SAME

This invention was made with government support under CA 57846 awarded by the National Institutes of Health. The government has certain rights in the invention.

BACKGROUND OF THE INVENTION

FIELD OF THE INVENTION

The present invention relates generally to tumor biology and cancer therapy and more specifically to a melanoma associated antigen and T cell epitopes of the antigen, as well as to methods of using such compositions to stimulate an immune response against melanoma cells.

BACKGROUND INFORMATION

The incidence of malignant melanoma has been increasing rapidly. Although ultraviolet radiation is believed to be the primary cause of melanoma, familial occurrence of the disease indicates that hereditary factors also may be involved. Unfortunately, methods for successfully treating melanoma have not kept pace with the increasing incidence.

In general, melanoma, like other cancers, is treated using surgical, chemotherapeutic or, in some cases, radiotherapeutic methods, or combinations of these methods. Surgical methods, however, can be curative only when the melanoma is detected early and has not metastasized. Similarly, radiotherapy, when used, generally only is effective when the tumor is localized. In the majority of cases, however, the melanoma has metastasized by the time it has been diagnosed and,

therefore, chemotherapy is indicated, sometimes in combination with surgery or radiotherapy. However, chemotherapy suffers from the disadvantage that it generally is not specific for the melanoma cells, but also kills rapidly dividing normal cells. In fact, toxicity to normal cells generally limits the dose of chemotherapy that a patient can tolerate. In addition, the cancer cells can become resistant to the chemotherapeutic agent and, therefore, refractory to the treatment. Thus, the duration of response to chemotherapy, radiotherapy or surgery can be too brief.

In theory, immunotherapy holds great promise for treating a cancer such as melanoma, particularly because it can be effective against disseminated disease and because it is expected to be specific only for the cancer cells. Efforts at immunotherapy of melanoma have been attempted using crude vaccines being composed, for example, of "killed" melanoma cells isolated either from the patient to be treated or from another patient, lysates of such cells or cell extracts. For example, a potential therapeutic melanoma vaccine, designated MELACINE, has been formulated from lysates of melanoma cells obtained from two different patients and has produced some positive results when used to treat patients having substantial disease or minimal residual disease. However, use of crude melanoma vaccines in immunotherapy is problematic, for example, because the precise antigenic composition of such vaccines is largely undefined. It is generally believed that more effective immunotherapy requires the identification and isolation of proteins that are expressed relatively specifically by melanoma cells, preferably on their surface, but are not expressed on normal cells. Thus, a need exists to identify melanoma associated antigens. The present invention satisfies this need and provides additional advantages.

SUMMARY OF THE INVENTION

The present invention provides a substantially purified polypeptide portion of a melanoma associated antigen, MG50, comprising the amino acid sequence shown as SEQ ID NO: 2, and provides substantially purified T cell epitopes, comprising a contiguous amino acid sequence of SEQ ID NO: 2, particularly a contiguous sequence within the sequence shown as amino acids 1187 to 1447 of SEQ ID NO: 2. For example, the invention provides cytotoxic T cell epitopes, comprising 8 to 11 contiguous amino acids of SEQ ID NO: 2, such as the cytotoxic T cell epitopes RPRPEQEPLP (SEQ ID NO: 4), DVTSGNTVY (SEQ ID NO: 5) and VLFCAWGTL (SEQ ID NO: 6), and provides helper T cell epitope comprising 12 to 25 contiguous amino acids of SEQ ID NO: 2.

In one embodiment, the invention provides MG50 cytotoxic T cell epitopes fused to a signal peptide or a functional portion thereof, which facilitates presentation of the epitope as a complex with an MHC molecule at the surface of an antigen presenting cell. For example, the invention provides MRYMILGLLLAALAAVCSARPRPEQEPLP (SEQ ID NO: 21), which is a cytotoxic T cell epitope fused to a signal peptide. In another embodiment, the invention provides a chimeric polypeptide, comprising an MG50 polypeptide encoded by SEQ ID NO: 1 or an MG50 T cell epitope encoded by SEQ ID NO: 1, fused to a second polypeptide, which is not MG50 or an MG50 T cell epitope, wherein the second polypeptide can facilitate detection of the MG50 component, for example, or can render an MG50 T cell epitope immunogenic.

The invention also provides antibodies, or antigen binding fragments thereof, that specifically bind the MG50 melanoma associated antigen (SEQ ID NO: 2) or a

peptide encoded by SEQ ID NO: 1, for example, an MG50 T cell epitope. If desired, an antibody of the invention can specifically bind an MG50 T cell epitope that is fused to a signal peptide or a functional portion thereof
5 or can specifically bind a chimeric polypeptide of the invention. Such antibodies, which can be monoclonal antibodies, are useful, for example, to prepare an anti-idiotypic antibody, which specifically binds to the antibody of the invention and provides a mimic of the
10 MG50 antigen used to raise the antibody.

The invention also provides a substantially purified nucleic acid molecule having the nucleotide sequence shown as SEQ ID NO: 1 (nucleotides 1 to 6848), including subsequences shown as nucleotides 1 to 5509,
15 nucleotides 1 to 3555, nucleotides 1 to 4336, nucleotides 3555 to 4336, nucleotides 3555 to 5509 and nucleotides 3555 to 6848. The invention further provides a substantially purified nucleic acid molecule encoding an MG50 polypeptide comprising SEQ ID NO: 2. In addition,
20 the invention provides nucleic acid molecules encoding MG50 T cell epitopes, such nucleic acid molecules comprising a portion of SEQ ID NO: 1, particularly of nucleotides 1 to 5509 or nucleotides 3555 to 4336 of SEQ ID NO: 1, or comprising a nucleotide sequence encoding a
25 portion of SEQ ID NO: 2, particular amino acids 1187 to 1447 of SEQ ID NO: 2. The invention further provides vectors containing a nucleic acid molecule of the invention, for example, expression vectors or viral vectors, and provides cells containing such vectors. In
30 an embodiment of the invention, antigen presenting cells, which contain and express a nucleic acid molecule of the invention, are provided, such cells which can present an MG50 T cell epitope complexed with an MHC molecule at its surface.

The present invention also provides methods of identifying the presence of an MG50 melanoma associated antigen in an individual, for example, by contacting a biological sample obtained from the individual with an antibody that specifically binds an MG50 antigen, wherein specific binding of the antibody to a component of the sample identifies the presence of the MG50 melanoma associated antigen in the individual. Conversely, the invention provides methods of identifying the presence in an immune response against an MG50 melanoma associated antigen in an individual, by contacting a biological sample obtained from the subject with a peptide encoded by SEQ ID NO: 1 and detecting an immunoeffector function of the sample due to contact with the peptide, thereby identifying the presence of an immune response against an MG50 melanoma associated antigen in the individual. For example, the ability of an MG50 peptide comprising an MG50 T cell epitope to stimulate the proliferation of T cells, which are a component of the biological sample, identifies the presence of an immune response in the individual from whom the sample was obtained.

The invention also provides methods for producing a population of antigen presenting cells that express an MG50 T cell epitope complexed with an MHC molecule on their surfaces. For example, the antigen presenting cells can be contacted with an MG50 melanoma associated antigen comprising the polypeptide encoded by SEQ ID NO: 1 or an MG50 T cell epitope encoded by SEQ ID NO: 1, which can be fused to a signal peptide or a functional portion thereof, if desired. In addition, the antigen presenting cells can be contacted with a nucleic acid molecule encoding an MG50 polypeptide, for example, SEQ ID NO: 1 or a nucleic acid molecule encoding SEQ ID NO: 2, or with a nucleic acid molecule encoding an MG50 T cell epitope, which can be fused to a signal peptide. Accordingly, the invention further provides populations

of antigen presenting cells produced by such a method of the invention and provides methods for stimulating T lymphocytes to react specifically against cancer cells expressing an MG50 melanoma associated antigen by
5 contacting the T cells with such antigen presenting cells.

The invention also provides methods for treating an individual having a cancer in which the cancer cells express an MG50 melanoma associated antigen.
10 For example, such an individual can be treated by administration of antigen presenting cells that express an MG50 T cell epitope complexed with an MHC molecule on their surfaces or can be treated by administration of T lymphocytes that have been stimulated *in vitro* to react
15 with cancer cells expressing an MG50 melanoma associated antigen. In addition, an individual can be treated by administration of a composition comprising an MG50 melanoma associated antigen, or a nucleic acid molecule encoding such an antigen, for example, an MG50
20 polypeptide encoded by SEQ ID NO: 1 or an MG50 T cell epitope encoded by SEQ ID NO: 1. If desired, the composition can contain an immunostimulatory agent such as an adjuvant, for example, DETOX, or a cytokine, for example, interleukin-2 or interferon- α , or the
25 immunostimulatory agent can be administered separately. The invention also provides methods of preventing the formation of a cancer due to cancer cells expressing an MG50 melanoma associated antigen.

BRIEF DESCRIPTION OF THE DRAWINGS

30 Figures 1A to 1E show the nucleotide sequence (SEQ ID NO: 1) encoding a portion of the MG50 melanoma associated antigen. The open reading frame is underlined. One potential polyadenylation signal is

double underlined and a second is indicated in bold and underlined.

Figure 2 shows the amino acid sequence (SEQ ID NO: 2) of the MG50 melanoma associated antigen encoded by the open reading frame shown in Figure 1 (SEQ ID NO: 1).

DETAILED DESCRIPTION OF THE INVENTION

The present invention provides a substantially purified polypeptide portion of a melanoma associated antigen, MG50, comprising the amino acid sequence shown as SEQ ID NO: 2. In addition, the invention provides substantially purified T cell epitopes, comprising a contiguous amino acid sequence of SEQ ID NO: 2 or an amino acid sequence encoded by SEQ ID NO: 1, particularly an amino acid sequence encoded by nucleotides 1 to 5509 of SEQ ID NO: 1, preferably by nucleotides 3555 to 4336 of SEQ ID NO: 1. As used herein, the term "substantially purified," when used in reference to an MG50 polypeptide, means that the polypeptide is relatively free from contaminating lipids, proteins, nucleic acids or other cellular material normally associated with an MG50 polypeptide in a cell. Methods for obtaining a substantially purified MG50 polypeptide of the invention are provided, below.

As disclosed herein, MG50 or a T cell epitope of MG50 is useful for stimulating specific reactivity of immunoeffector cells against cells expressing the MG50 melanoma associated antigen. MG50 is considered a melanoma associated antigen because it is expressed on melanoma cells obtained from different individuals. In addition, MG50 is considered a shared tumor antigen because it is expressed on different types of cancer cells, including, for example, melanoma cells, lung cancer cells and rhabdomyosarcoma cells.

Various melanoma associated antigens, which are shared among melanoma cells in different patients, have been identified. The product of the MAGE-1 gene, the MZ2-E antigen (hereinafter "MAGE-1"), was the first shared melanoma antigen identified (Treversari et al., Immunogenetics 35:145- (1991); van der Bruggen et al., Science 254:1643- (1991)). MAGE-1 was determined to stimulate cytotoxic T cell ("Tc cell") activity and is HLA-A1 restricted (Traversi et al., J. Exp. Med. 176:1453-1457 (1992)). Additional related melanoma antigens subsequently were identified and named MAGE-2, MAGE-3 and MAGE-4 based on their homology to MAGE-1. The MAGE antigens do not appear to be expressed in normal tissues.

MART-1 is another example of a shared human melanoma antigen (Kawakami et al., Proc. Natl. Acad. Sci., USA 91:3515-3519 (1994)). MART-1 is expressed in melanoma cells and, to a lesser extent, melanocytes and retina, and, therefore, appears restricted to cells of melanocyte lineage. Unlike the MAGE antigens, which are HLA-A1 restricted, MART-1 is presented in association with HLA-A2 molecules, which are expressed in about 50% of the population. In comparison, only about 10% of the population express HLA-A1 molecules.

MG50 also is a melanoma associated antigen and was cloned from the M1 melanoma cell line by subtractive hybridization; the gene encoding MG50 is present on chromosome 2 (Hutchins et al., Cancer Res. 51:1418-1425 (1991); Weiler, "Molecular Characterization of a Novel Human Melanoma Associated Gene (MG50), Dissertation submitted to the University of Southern California, December, 1993; Weiler et al., Genomics 22:243-244 (1994); Genome Data Bank, Accession No.: locus D2S448 (G00-252-144), each of which is incorporated herein by reference). As disclosed herein, the MG50 gene, which

encodes the polypeptide shown as SEQ ID NO: 2, contains cryptic coding sequences located downstream of a polyadenylation signal.

A nucleic acid molecule encoding MG50 was
5 cloned and, early in the sequencing of the cDNA, a peptide, CSEQPFPEHTASVQHAD (SEQ ID NO: 3) was prepared based on the MG50 cDNA sequence (see Weiler, *supra*, 1993; referred to a "pep-50"). This peptide, which was suspected of being encoded by a nucleic acid sequence
10 located downstream of the MG50 coding sequence, stimulated proliferation of melanoma specific T cells (Weiler, *supra*, 1993). As disclosed herein, CSEQPFPEHTASVQHAD (SEQ ID NO: 3) is encoded by a cryptic MG50 coding sequence, as is the peptide RPRPEQEPLP (SEQ
15 ID NO: 4), which also stimulates proliferation of melanoma specific T cells. Remarkably, RPRPEQEPLP (SEQ ID NO: 4) and CSEQPFPEHTASVQHAD (SEQ ID NO: 3) are not encoded by the same reading frame of SEQ ID NO: 1. Specifically, while RPRPEQEPLP (SEQ ID NO: 4) is encoded
20 by the same reading frame as the remainder of the MG50 polypeptide, although it is downstream of a potential polyadenylation signal, CSEQPFPEHTASVQHAD (SEQ ID NO: 3) is out of frame with respect to the MG50 coding sequence. Thus, T cell epitopes of MG50 can be identified in
25 cryptic coding regions of SEQ ID NO: 1 and can comprise an amino acid sequence encoding by various reading frames of the cryptic coding sequence.

RPRPEQEPLP (SEQ ID NO: 4) is a T cell epitope that is recognized by cytotoxic T cells in the context of
30 the Class I MHC molecule HLA-B7. Additional MG50 peptide sequences having the characteristics of T cell epitopes that are recognized by HLA-A1 MHC molecules (SEQ ID NO: 5) or by HLA-A2 MHC molecules (SEQ ID NOS: 6-17) also have been identified (see Table 1). Such MG50 epitopes
35 were identified by homology to consensus HLA-A1 or HLA-A2

epitopes (see Kaat et al., J. Immunol. 152:3904-3912 (1994); Falk and Rotzschko, Sem. Immunol. 5:81-94 (1993), each of which is incorporated herein by reference). The availability of MG50 T cell epitopes that are presented
5 in the context of HLA-A2 MHC molecules provides that advantage that HLA-A2 molecules are expressed by a relatively large number of individuals as compared to HLA-B7.

The invention provides cytotoxic T cell
10 epitopes, comprising 8 to 11 contiguous amino acids encoded by SEQ ID NO: 1, for example, the cytotoxic T cell epitopes RPRPEQEPLP (SEQ ID NO: 4), DVTSGNTVY (SEQ ID NO: 5) and VLFCAWGTL (SEQ ID NO: 6). As used herein, the term "cytotoxic T cell epitope" means a T cell
15 epitope that is recognized by and stimulates cytotoxic T cells. Also provided are helper T cell epitopes, comprising 12 to 25 contiguous amino acids encoded by SEQ ID NO: 1. As used herein, the term "helper T cell epitope" means a T cell epitope that is recognized by and
20 stimulates helper T cells.

As used herein, the term "T cell epitope" means a peptide that is complexed with an MHC molecule and, when complexed with the MHC molecule, can be bound to a T cell receptor. In addition, the term "T cell epitopic
25 fragment" is used herein to mean a peptide portion of a protein, which can be formed due to proteolysis of the protein, that has the characteristics of a T cell epitope as defined above. In view of these definitions, it should be recognized that the terms "T cell epitope" and
30 "T cell epitopic fragment" often can be used synonymously. However, while a "T cell epitopic fragment" specifically comprises a peptide sequence that is present in a protein, a "T cell epitope" can comprise a peptide containing an amino acid sequence that is the
35 same as or different from the corresponding sequence

present in the protein from which the epitope was derived. Thus, the term "T cell epitope" broadly encompasses a T cell epitopic fragment. In addition, the term "peptide" or "peptide portion," when used in reference to MG50, means an amino acid sequence of at least two contiguous amino acids of SEQ ID NO: 2 (amino acids 1 to 1497), particularly of amino acids 1187 to 1447 of SEQ ID NO: 2, and that are unique to MG50.

A T cell epitope varies in size based on the MHC molecule that binds the epitope. Specifically, class I MHC molecules bind peptides containing about 8 to 11 amino acids, generally peptides containing 8 to 10 amino acids and, most often, 9 or 10 amino acids. In comparison, class II MHC molecules bind peptides containing about 12 to 25 amino acids, generally peptides containing 13 to 18 amino acids.

Class I MHC molecules are, or can be, expressed by all nucleated cells, including antigen presenting cells (see below), and present T cell epitopes to Tc cells. The epitopes presented by class I MHC molecules often are produced by proteolysis of endogenously expressed proteins, including proteins expressed in virally infected cells and in tumor cells. The epitope likely associates with the class I molecule in the endoplasmic reticulum, then the complex is transported to the cell surface. Tc cells, which express the CD8 surface antigen ("CD8+") and a T cell receptor, then bind the epitope associated with the class I molecule, thereby activating the effector function of the Tc cells (see, generally, Kuby, "Immunology" 3d ed. (W.H. Freeman and Co., 1997)).

In comparison to class I molecules, which are expressed on by nucleated cells, class II MHC molecules only are expressed by antigen presenting cells (APC's),

including B lymphocytes ("B cells"), dendritic cells, mononuclear phagocytic cells, macrophages, including Langerhans cells and, in humans, venular endothelial cells, and present a T cell epitope to helper T cells ("Th cells"), stimulating the Th cells, such stimulation being effective in immunity to tumors (see, generally, Abbas et al., "Cellular and Molecular Immunology," 2d ed. (W.B. Saunders Co. 1995); Jones and Mitchell, Trends Biotechnol. 14:349-355 (1996), each of which is incorporated herein by reference; see, also, Kuby, *supra*, 1997). Thus, APC's express both class I and class II MHC molecules and, therefore, can activate Tc cells and Th cells.

The epitopes that are bound by class II molecules generally are derived by proteolysis of exogenous proteins, which are internalized in the APC by phagocytosis or endocytosis. In addition, APC's, such as macrophages, can express co-stimulatory B7 molecules, B7-1 (CD80) and B7-2 (CD86), which are recognized by a cell surface molecule (CD28) that is expressed by certain T cells, including naive T cells, and is involved in activation of the T cells. Binding of a T cell epitope and B7 molecule by Th cells stimulates activation of two subsets of Th cells, Th1 cells, which express interleukin-2 (IL-2), interferon- γ , tumor necrosis factor- β and tumor necrosis factor- α and are involved in the cell-mediated immune functions, including activation of Tc cells; and Th2 cells, which secrete IL-4, IL-5, IL-6 and IL-10 and are involved in the activation of B cells (Quan and Mitchell, in "Current Research and Clinical Management of Melanoma" (Kluwer Academic Publ. 1993), which is incorporated herein by reference; see pages 257-277; see, also, Kuby, *supra*, 1997; chaps. 1, 10 and 12; Jones and Mitchell, *supra*, 1996).

In an embodiment of the invention, MG50 cytotoxic T cell epitopes fused to a signal peptide or a functional portion thereof are provided. Fusion of a signal peptide, for example, to the N-terminus of a cytotoxic MG50 T cell epitope, can facilitate presentation of the epitope as a complex with an MHC molecule at the surface of an antigen presenting cell (see Minev et al., Cancer Res. 54:4155-4161 (1994), which is incorporated herein by reference). For example, the invention provides MRYMILGLLLALAAVCSARPRPEQEPLP (SEQ ID NO: 21) and MTNKCLLOIALLLCFSTTALSRPRPEQEPLP (SEQ ID NO: 22), which contain the cytotoxic T cell epitope, RPRPEQEPLP (SEQ ID NO: 4), fused to two different signal peptides (signal peptide is underlined).

Signal peptides are well known in the art and consist generally of three functional portions: a basic N-terminal region of about 1 to 3 positively charged amino acids; a central hydrophobic region of about 8 to 12 hydrophobic amino acids; and a polar C-terminal region of about 5 to 7 amino acids with higher average polarity than the central hydrophobic region. Recognition of a signal peptide by a signal peptidase, which is located within the endoplasmic reticulum of a eukaryotic cell, results in cleavage of the signal peptide from the remainder of the molecule, for example, the MG50 T cell epitope.

For use in the present invention, a signal peptide or a functional portion thereof can be based on any naturally occurring signal sequence or can be a non-naturally occurring sequence having the general characteristics of a signal peptide or of a functional portion of the signal peptide. As used herein, the term "functional portion," when used in reference to a signal peptide, means the basic N-terminal region of about 1 to 3 positively charged amino acids; the central hydrophobic

region of about 8 to 12 hydrophobic amino acids; or polar C-terminal region of about 5 to 7 amino acids with higher average polarity than the central hydrophobic region. As disclosed herein, the substitution, for example, of the central hydrophobic 8 to 12 amino acids of a signal peptide with an MG50 T cell epitope having the appropriate hydrophobicity, provides a unique type of T cell epitope having characteristics of an MG50 T cell epitope fused to a signal sequence. Contact of an APC with such an MG50 T cell epitope results in efficient transport of the epitope complexed with an MHC molecule to the surface of an APC.

In another embodiment, the invention provides a chimeric MG50 polypeptide, comprising an MG50 polypeptide, comprising SEQ ID NO: 2, or a peptide portion of MG50 such as an MG50 T cell epitope encoded by SEQ ID NO: 1, fused to a second peptide or polypeptide, which is not MG50 or a peptide portion of MG50. A chimeric MG50 polypeptide provides certain advantages. For example, where the second polypeptide is an enzyme, such as alkaline phosphatase, horseradish peroxidase or luciferase, detection of the MG50 component of the fusion is facilitated by detecting the presence of the enzyme activity. Other such detectable markers, for example, a FLAG epitope, also can be fused to an MG50 T cell epitope or, if desired, to an MG50 polypeptide, for the purpose of detecting the presence of the MG50 or MG50 T cell epitope. Such a detectably labeled chimeric polypeptide is useful, for example, in an immunoassay to identify the presence of anti-MG50 antibodies or of MG50 reactive immunoeffector cells in a biological sample obtained from a subject.

A chimeric polypeptide of the invention also can be MG50 or an MG50 T cell epitope fused to a second polypeptide such as glutathione-S-transferase (GST) or

the His-6 peptide. Such a chimeric polypeptide can be particularly useful for purifying the MG50 or MG50 T cell epitope. For example, GST readily binds to glutathione, which can be attached to an insoluble matrix, thereby
5 providing a simple affinity chromatography method of purifying a GST-MG50 chimeric polypeptide. Similarly, the His-6 sequence readily binds to a cation such as nickel ion, thus allowing for purification of a His-6-MG50 chimeric polypeptide. Such chimeric
10 polypeptides also can be used to purify antibodies that specifically bind to an MG50 component of the chimeric polypeptide.

A chimeric polypeptide of the invention also can be a peptide portion of the MG50 polypeptide, for
15 example, an MG50 T cell epitope, fused to a carrier protein such as bovine serum albumin, bovine gamma-globulin, human gamma-globulin, keyhole limpet hemocyanin, or ovalbumin. Such a chimeric polypeptide can render a haptenic MG50 peptide immunogenic and,
20 therefore, can be useful for inducing an anti-MG50 antibody response, thus providing a means for obtaining anti-MG50 antibodies (see Harlow and Lane, "Antibodies: A laboratory manual" (Cold Spring Harbor Laboratory Press 1988), which is incorporated herein by reference; see,
25 also, Kuby, *supra*, 1997).

MG50, an MG50 T cell epitope, an MG50 T cell epitope fused to a signal sequence, or a chimeric MG50 polypeptide can be produced by a variety of routine methods, including, as appropriate, biochemical
30 purification, recombinant DNA methods or chemical synthesis. An MG50 polypeptide or MG50 T cell epitope, for example, can be produced by recombinant DNA methods (see, generally, Sambrook et al., Sambrook et al., Molecular Cloning: A Laboratory Manual (Cold Spring
35 Harbor Laboratory Press, 1989), which is incorporated

herein by reference). For example, an MG50 T cell epitope can be produced by cloning a nucleic acid encoding the epitope into an expression vector such as a baculovirus vector, then isolating the expressed epitope from an appropriate insect host cell. In addition, an MG50 polypeptide also can be expressed in a mammalian cell, where it can be post-translationally modified in a manner expected for a native MG50 protein. Appropriate expression vectors and host cells are well known in the art (see Kriegler, Gene Transfer and Expression. A Laboratory Manual, W.H. Freeman and Co., 1991, which is incorporated herein by reference) and are commercially available.

An MG50 T cell epitope, for example, can be synthesized using well known chemical methods, including, for example, automated solid phase methods. Chemical synthesis of an MG50 T cell epitope can be particularly desirable because the method allows for the introduction of amino acid analogs such as (D)-amino acids into the peptide, if desired. The incorporation of a (D)-amino acid, for example, can increase the stability of the T cell epitope, which can be particularly useful for preparing a vaccine or for preparing a diagnostic kit.

Native MG50 protein can be purified from a melanoma cell lysate using, for example, an antibody of the invention. In addition, MG50 T cell epitopes can be purified from antigen presenting cells obtained from an individual suffering from a cancer that expresses MG50. Methods for obtaining such T cell epitopes are well known in the art and include immunoaffinity chromatography, gel filtration chromatography or gel electrophoresis (see, for example, Hunt et al., Science 256:1817-1820 (1992); Chicz et al., Nature 358:764-768 (1992); see, also, Deutscher, Guide to Protein Purification (Academic Press,

Inc. 1990), each of which is incorporated herein by reference.

The invention also provides antibodies, or antigen binding fragments thereof, that specifically bind to an MG50 melanoma associated antigen comprising SEQ ID NO: 2, particularly to amino acids 1187 to 1447 of SEQ ID NO: 2, or to a peptide encoded by SEQ ID NO: 1, for example, an MG50 T cell epitope. If desired, an antibody of the invention can specifically bind an MG50 T cell epitope that is fused to a signal peptide or a functional portion thereof or can specifically bind a chimeric polypeptide of the invention. Such antibodies, which can be monoclonal antibodies, are useful, for example, to prepare an anti-idiotypic antibody, which specifically binds to the antibody of the invention and provides a mimic of the MG50 antigen used to raise the antibody. The antibody also can be labeled with a detectable label, for example, a radionuclide, biotin, or an enzyme, using known methods (see, Harlow and Lane, *supra*, 1988; Hermanson, "Bioconjugate Techniques" (Academic Press 1996), which is incorporated herein by reference).

As used herein, the term "antibody" is used in its broadest sense to include polyclonal and monoclonal antibodies, as well as antigen binding fragments of such antibodies. An antibody of the invention, or an antigen binding fragment thereof, is characterized in that it specifically can bind with an MG50 epitope with an affinity of at least about $1 \times 10^5 \text{ M}^{-1}$ and, generally, at least about $1 \times 10^6 \text{ M}^{-1}$. Such antigen binding fragments of an antibody include, for example, Fab, F(ab')_2 , Fd and Fv fragments that retain specific binding activity for an MG50 epitope.

An antibody of the invention can be either naturally occurring or non-naturally occurring, and can

include, for example, a single chain antibody, a chimeric, bifunctional and humanized antibody, as well as an antigen-binding fragment thereof. A non-naturally occurring antibody can be constructed using solid phase peptide synthesis, can be produced recombinantly or can be obtained, for example, by screening combinatorial libraries consisting of variable heavy chains and variable light chains (see Huse et al., Science 246:1275-1281 (1989), which is incorporated herein by reference). These and other methods of making, for example, a chimeric, humanized, CDR-grafted, single chain, or bifunctional antibody are well known to those skilled in the art (Hoogenboom et al., U.S. Patent No. 5,564,332, issued October 15, 1996; Winter and Harris, Immunol. Today 14:243-246 (1993); Ward et al., Nature 341:544-546 (1989); Harlow and Lane, *supra*, 1988); Hilyard et al., Protein Engineering: A practical approach (IRL Press 1992); Borrabeck, Antibody Engineering, 2d ed. (Oxford University Press 1995); each of which is incorporated herein by reference).

Methods for raising polyclonal antibodies, for example, in a rabbit, goat, mouse or other mammal, are well known in the art. In addition, monoclonal antibodies can be obtained using methods that are well known and routine in the art (Harlow and Lane, *supra*, 1988). Essentially, spleen cells from a mouse immunized with MG50 or a peptide portion thereof can be fused to an appropriate myeloma cell line such as SP/02 myeloma cells to produce hybridoma cells. Cloned hybridoma cell lines can be screened using labeled antigen to identify clones that secrete the desired monoclonal antibodies. Hybridomas expressing, for example, anti-MG50 monoclonal antibodies having a desirable specificity and affinity can be isolated and utilized as a continuous source of the antibodies, which are useful, for example, for preparing standardized kits containing the antibody.

Similarly, a recombinant phage that expresses, for example, a single chain anti-MG50 antibody provides a monoclonal antibody that can be used for preparing a kit.

A monoclonal antibody specific for MG50 or a peptide portion of MG50 can be used to prepare anti-idiotypic antibodies, which present an epitope that mimics the epitope recognized by the monoclonal antibody used to prepare the anti-idiotypic antibodies. Where the epitope to which the monoclonal antibody includes, for example, an MG50 T cell epitope, the anti-idiotypic antibody can be useful for detecting the presence of MG50 reactive T cells in a biological sample obtained from an individual. In addition, vaccines containing anti-idiotypic antibodies can have antitumor prophylactic effects and induce the involvement of Th cells (Mitchell, Brit. Med. Bull. 51:631-646 (1995); Quan et al., J. Clin. Oncol. 15:2103-2110 (1997), each of which is incorporated herein by reference).

The invention also provides a substantially purified nucleic acid molecule having the nucleotide sequence shown as SEQ ID NO: 1 (see, also, Weiler, *supra*, 1993). In particular, the invention provides subsequences of SEQ ID NO: 1, including nucleotides 1 to 5509, nucleotides 1 to 3555, nucleotides 1 to 4336, nucleotides 3555 to 4336, nucleotides 3555 to 5509 and nucleotides 3555 to 6848. The portion of SEQ ID NO: 1 shown as nucleotides 1 to 5510 also is available at GenBank Accession No. D86983 (submitted by N. Nomura; August 2, 1996). As used herein, the term "substantially purified," when used in reference to a nucleic acid molecule of the invention, means that the nucleic acid molecule is relatively free from contaminating lipids, proteins, nucleic acids or other cellular material normally associated with a nucleic acid molecule in a cell. A substantially purified nucleic acid molecule of

the invention can be obtained by chemical synthesis of the nucleotide sequence shown as SEQ ID NO: 1 or by cloning the molecule using, for example, a method of the polymerase chain reaction (PCR), wherein appropriate
5 primers are selected based on SEQ ID NO: 1.

Due to the degeneracy of the genetic code and in view of the disclosed MG50 amino acid sequence shown in SEQ ID NO: 2, particularly of amino acids 1187 to 1447 of SEQ ID NO: 2, additional nucleic acid molecules of the
10 invention would be well known to those skilled in the art. Such nucleic acid molecules have a nucleotide sequence that is different from the sequence shown as nucleotides 1 to 4488, particularly of nucleotides 3555 to 4336, of SEQ ID NO: 1 but, nevertheless, encode the
15 amino acid sequence shown as amino acids 1 to 1497, particularly amino acids 1187 to 1447, respectively, of SEQ ID NO: 2. Thus, the invention provides a nucleic acid molecule comprising a nucleotide sequence encoding an MG50 polypeptide comprising SEQ ID NO: 2. In
20 addition, the invention provides nucleic acid molecules encoding MG50 T cell epitopes, such nucleic acid molecules comprising a portion of SEQ ID NO: 1 or comprising a nucleotide sequence encoding a portion of SEQ ID NO: 2.

25 As used herein, the term "a nucleic acid molecule encoding," when used in reference to MG50 or to a peptide portion of MG50, including an MG50 T cell epitope, indicates 1) the polynucleotide sequence of one strand of a double stranded DNA molecule comprising the
30 nucleotide sequence that codes for MG50 or a peptide portion of MG50 and can be transcribed into an RNA that encodes MG50 or the peptide, or 2) an RNA molecule, which can be translated into MG50 or a peptide portion thereof. It is recognized that a double stranded DNA molecule also
35 comprises a second polynucleotide strand that is

complementary to the coding strand and that the disclosure of a polynucleotide sequence comprising a coding sequence necessarily discloses the complementary polynucleotide sequence. Accordingly, the invention
5 provides polynucleotide sequences, including, for example, polydeoxyribonucleotide or polyribonucleotide sequences that are complementary to the nucleotide sequence shown as SEQ ID NO: 1 or to a nucleic acid molecule encoding MG50, comprising the amino acid
10 sequence shown as SEQ ID NO: 2, or to a peptide portion of SEQ ID NO: 2. As used herein, the term "polynucleotide" is used in its broadest sense to mean two or more nucleotides or nucleotide analogs linked by a covalent bond.

15 The invention also provides nucleotide sequences of SEQ ID NO: 1, particularly of nucleotides 3555 to 4336 of SEQ ID NO: 1, which specifically hybridize to a nucleic acid molecule encoding MG50. It is recognized, for example, that SEQ ID NO: 1 shares
20 regions of homology with nucleic acid molecules encoding mammalian peroxidases and Drosophila peroxidase. Thus, a nucleotide sequence of SEQ ID NO: 1, which is considered to be within the claimed invention, hybridizes under stringent hybridization conditions to a nucleotide
25 sequence encoding MG50, but not to a nucleic acid molecule encoding a mammalian peroxidase, such as those disclosed as GenBank Accession Numbers X15313, X15378, M29907, X14346, L77979, or the like, or to a nucleotide sequence encoding Drosophila peroxidase, such as that
30 disclosed as GenBank Accession No. U11052.

A nucleotide sequence of the invention is useful, for example, as a probe, which can hybridize to a nucleic acid molecule encoding MG50 and allow the identification of the nucleic acid molecule in a sample.
35 A nucleotide sequence of the invention is characterized,

in part, in that it is at least nine nucleotides in length, such sequences being particularly useful as primers for PCR, and can be at least fourteen nucleotides in length or, if desired, at least seventeen nucleotides in length, such nucleotide sequences being particularly useful as hybridization probes, although such sequences also can be used for PCR. In addition, a nucleotide sequence of the invention comprises at least six nucleotides, preferably at least nine nucleotides, 5' to nucleotide 5509 of SEQ ID NO: 1, where SEQ ID NO: 1 is shown in the conventional manner from the 5'-terminus (Figure 1A; upper left) to the 3'-terminus, most preferably at least nine contained within nucleotides 3555 to 4336 of SEQ ID NO: 1.

The invention also provides vectors comprising a nucleic acid molecule of the invention and host cells, which are appropriate for maintaining such vectors. Vectors, which can be cloning vectors or expression vectors, are well known in the art and commercially available. An expression vector comprising a nucleic acid molecule of the invention, which can encode, for example, MG50 or a T cell epitope thereof, can be used to express the nucleic acid molecule in a cell.

In general, an expression vector contains the elements necessary to achieve, for example, sustained transcription of the nucleic acid molecule, although such elements also can be inherent to the nucleic acid molecule cloned into the vector. In particular, an expression vector contains or encodes a promoter sequence, which can provide constitutive or, if desired, inducible expression of a cloned nucleic acid sequence, a poly-A recognition sequence, and a ribosome recognition site, and can contain other regulatory elements such as an enhancer, which can be tissue specific. The vector also contains elements required for replication in a

procaryotic or eukaryotic host system or both, as desired. Such vectors, which include plasmid vectors and viral vectors such as bacteriophage, baculovirus, retrovirus, lentivirus, adenovirus, vaccinia virus, semliki forest virus and adeno-associated virus vectors, are well known and can be purchased from a commercial source (Promega, Madison WI; Stratagene, La Jolla CA; GIBCO/BRL, Gaithersburg MD) or can be constructed by one skilled in the art (see, for example, Meth. Enzymol., Vol. 185, D.V. Goeddel, ed. (Academic Press, Inc., 1990); Jolly, Canc. Gene Ther. 1:51-64 (1994); Flotte, J. Bioenerg. Biomemb. 25:37-42 (1993); Kirshenbaum et al., J. Clin. Invest 92:381-387 (1993), which is incorporated herein by reference).

A nucleic acid molecule, either alone or contained a vector, can be introduced into a cell by any of a variety of methods known in the art (see Ausubel et al., Current Protocols in Molecular Biology, John Wiley and Sons, Baltimore, MD (1994); Chang, Somatic Gene Therapy, Chap. 11 (CRC Press, Inc., 1995), each of which is incorporated herein by reference; see, also, Sambrook et al., *supra*, 1989). Such methods include, for example, transfection, lipofection, microinjection, electroporation and infection with recombinant viral vectors or the use of liposomes. Introduction of a nucleic acid molecule by infection with a viral vector is particularly advantageous in that it can efficiently introduce the nucleic acid molecule into a cell *ex vivo* or *in vivo* (see, for example, U.S. Patent No. 5,399,346, issued March 21, 1995, which is incorporated herein by reference).

The invention also provides methods for producing a population of antigen presenting cells (APC's), which can express an MG50 T cell epitope complexed with an MHC molecule on their surfaces. APC's

are well known in the art and include dendritic cells, mononuclear phagocytic cells, B lymphocytes, Langerhans cells or human venular cells. In one embodiment of the invention, APC's that contain and express a nucleic acid molecule of the invention are provided. Such a nucleic acid molecule can be introduced into an APC using methods as discussed above. In another embodiment, the APC's are contacted with an MG50 melanoma associated antigen encoded by SEQ ID NO: 1 or a peptide portion thereof, particularly a peptide portion encoded by a nucleotide sequence contained within nucleotides 1 to 5509 or within nucleotides 3555 to 4226 of SEQ ID NO: 1, or are contacted with an MG50 T cell epitope encoded by SEQ ID NO: 1, particularly by a sequence within nucleotides 1 to 5509 or within nucleotides 3555 to 4336 of SEQ ID NO: 1, which can be fused to a signal peptide or a functional portion thereof, if desired. Accordingly, the invention also provides populations of APC's that are produced by a method of the invention and express on their cell surfaces an MG50 T cell epitope complexed with an MHC molecule.

APC's can be contacted, for example, with an MG50 T cell epitope fused to a signal peptide to produce a population of APC's encompassed within the claimed invention. The MG50 T cell epitope is loaded into the cytosol of T cells using osmotic lysis of pinocytic vesicles. T cells exposed to hypertonic medium take-up the fusion peptides due to the formation of pinocytic vesicles in the medium. The pinocytic vesicles break in the cytosol when the cells are placed in hypotonic culture medium, due to the increased internal osmotic pressure. The signal sequence then facilitates translocation of the MG50 T cell epitope from the cytosol into the endoplasmic reticulum, thereby increasing the efficiency with which the epitope is presented at the cell surface complexed with an MHC molecule.

APC's produced by a method of the invention can present an MG50 T cell epitope with a class II molecule or a co-stimulatory B7 molecule to a T cell to activate the T cell. Thus, the invention further provides methods
5 for stimulating T lymphocytes to react specifically against cancer cells expressing an MG50 melanoma associated antigen by contacting the T cells with an APC that presents an MG50 T cell epitope complexed with an MHC molecule on its surface. Although such a stimulation
10 can occur *in vivo*, for example, by administration of the APC's of the invention to an individual, such stimulation of T cells also can be performed *in vitro*. Accordingly, the invention provides an isolated population of T cells, which are specifically reactive with cancer cells that
15 express an MG50 melanoma associated antigen. In addition, it should be recognized that such a population of specifically reactive T cells can be obtained by contacting naive APC's and T cells *in vitro* with MG50 or an MG50 epitope, then isolating the T cells from the
20 APC's. Such *in vitro* methods of producing APC's that express an MG50 T cell epitope complexed with an MHC molecule on its cell surface or of producing T cells specifically reactive with a cell expressing an MG50 melanoma associated antigen are particularly useful
25 because the respected populations of the cells can be expanded such that a large number of the cells can be isolated. Furthermore, the skilled artisan will recognize that the APC's and the T cells can be autologous with respect to each other or can be
30 allogeneic (see, for example, Mitchell, *supra*, 1995).

The invention also provides methods for treating an individual having a cancer containing cancer cells that express an MG50 melanoma associated antigen. It is recognized that the methods of the invention can be
35 curative in some cases. However, a method of the

invention also can be useful where it is palliative and, therefore, increases the quality of life of an individual. In particular, the artisan skilled in cancer therapy will recognize that a method of the invention can
5 be particularly useful in combination with conventional cancer therapeutic modalities, including surgery, radiotherapy and chemotherapy.

An individual having a cancer containing cancer cells that express MG50 can be treated, for example, by
10 administration of APC's that express an MG50 T cell epitope complexed with an MHC molecule on their surfaces. Administration of such APC's can stimulate an active immune response in the subject by presenting MG50 T cell epitopes to the individual's T lymphocytes. In addition,
15 the individual can be treated by administration of T lymphocytes that have been stimulated *in vitro* to react with cancer cells expressing an MG50 melanoma associated antigen, thus providing a means of passive immunization of the individual.

20 The present invention also provides methods of treating an individual having a cancer containing cancer cells expressing an MG50 melanoma associated antigen by administering an MG50 vaccine to the individual. As used herein, the term "vaccine," when used in reference to the
25 present invention, means a formulation that is suitable for administration to a mammal, particularly a human, and contains an MG50 component selected from 1) an MG50 polypeptide, comprising SEQ ID NO: 2; 2) an MG50 T cell epitope encoded by SEQ ID NO: 1, which can be fused to a
30 signal peptide, if desired; 3) an anti-idiotypic antibody of the invention, which is a mimic of an MG50 epitope; or 4) a nucleic acid molecule encoding an MG50 polypeptide or MG50 T cell epitope.

A vaccine of the invention generally contains a pharmaceutically acceptable carrier, for example, an aqueous solution such as physiologically buffered saline or other solvent or vehicle such as a glycol, glycerol, oil such as olive oil or injectable organic ester. A pharmaceutically acceptable carrier also can include a physiologically acceptable compound that acts, for example, to stabilize the MG50 component of the formulation or to increase the absorption of the MG50 component. Physiologically acceptable compounds include, for example, carbohydrates, such as glucose, sucrose or dextrans, antioxidants, such as ascorbic acid or glutathione, chelating agents, low molecular weight proteins or other stabilizers or excipients.

If desired, a vaccine of the invention can contain an immunostimulatory agent such as an adjuvant, for example, DETOX (Ribi Immunochem), alum or Freund's complete or incomplete adjuvant. In addition, a vaccine can contain an immunostimulatory agent such as a cytokine, for example, interleukin-2 (IL-2), IL-4, IL-7, IL-12, IL-15, interferon- α (Ifn- α), Ifn- γ , granulocyte-macrophage colony stimulating factor (R&D Systems, Inc.; Minneapolis MN), an accessory molecule such as ICAM-1 or B7, or an agent from plants such as QS-21 (see Mitchell, *supra*, 1995; Jones & Mitchell, *supra*, 1996), although such agents also can be added separately from the vaccine, if desired. Other such agents include vectors that are rendered nonpathogenic, for example, by attenuation, liposomes and cell-sized microspheres (see Jones and Mitchell, *supra*, 1996). The skilled artisan will know how to formulate a vaccine of the invention using methods routine in the art. For example, a vaccine that includes about 10 μ g to about 10 mg of MG50 or an MG50 T cell epitope can be administered in conjunction with about 2.5 million units/meter² to about 20 million units/meter² of interferon- α . Such a combined modality

can be administered about three to five times per week and can be continued for up to about two years.

An MG50 vaccine can be administered for preventive purposes or for therapeutic purposes. The vaccine can be administered for preventive purposes, for example, to minimize the likelihood that a cancer such as melanoma, which expresses MG50, will occur in those individuals that are at high risk for the disease. Such individual include, for example, those suffering from familial dysplastic nevus syndrome or from atypical Spitz nevi, those individuals having a large number of moles or having an irregularly shaped mole, or those individuals living in high incidence regions such as Australia, Hawaii or the southwestern United States. In addition, an MG50 vaccine can be administered for therapeutic purposes to an individual suffering from a cancer that contains cancer cells expressing MG50 and can prevent the further growth or spread of the cancer or induce regression of the cancer. Such cancers can be, for example, melanoma, lung cancer or rhabdomyosarcoma and other cancers expressing MG50 can be identified using methods as disclosed herein.

An MG50 vaccine can be administered in a manner similar to other vaccines, for example, subcutaneously, orally, intradermally, intramuscularly or intravenously. In addition, following a first administration of the vaccine, it can be advantageous to administer one or more booster vaccinations. The need to administer a booster vaccination and the timing of such vaccinations can be determined experimentally by measuring, for example, the presence or proliferation of MG50 reactive Tc cells in the individual.

The MG50 component of a vaccine is administered to the individual in an amount that is sufficient to

stimulate an immune response, particularly a cellular immune response. Such an amount will vary, for example, depending on whether the MG50 component is an MG50 polypeptide or an MG50 T cell epitope or a nucleic acid molecule encoding the MG50 component. In addition, the amount will vary, for example, depending on whether stimulation of the immune response is *in vivo* or *in vitro*; whether the administration is a first administration or a booster administration; whether an immunostimulatory agent such as an adjuvant is administered; and, when administered *in vivo*, on the route of administration. In general, about 10 μ g to about 10 mg of an MG50 polypeptide or MG50 T cell epitope is administered per immunization. Methods for determining a sufficient amount of an MG50 component is required to stimulate an immune response are well within the means of the skilled artisan and generally are determined in Phase I and Phase II clinical trials (see, for example, Powell and Newman, Vaccine Design: The subunit and adjuvant approach (Plenum Publ. Corp.; 1994), which is incorporated herein by reference).

Where administration is of a nucleic acid molecule encoding MG50 or an MG50 T cell epitope, the nucleic acid molecule can be contained in a vector (see Goeddel, *supra*, 1990). A nucleic acid molecule can be inserted into such a vector using known methods (see, Sambrook, *supra*, 1989). A variety of vectors, including expression vectors, are available and contain, for example, a promoter such as the cytomegalovirus or SV40 promoter, which can direct expression of MG50 or an MG50 T cell epitope in a cell (see, Gacesa and Ramji, Vectors, Essential Data, John Wiley and Sons, NY (1994), which is incorporated herein by reference). Viral vectors based, for example, on a retrovirus, an adenovirus, an adeno-associated virus, a vaccinia virus, or the like,

are particularly useful (see, for example, Anderson et al., U.S. Patent No. 5,399,347, issued March 21, 1995; Lee et al., U.S. Patent No. 5,532,220, issued July 2, 1996; Collins et al., U.S. Patent No. 5,240,846, issued
5 August 31, 1993; and Ram et al., Cancer Res. 53:83-88 (1993); Karlsson et al., EMBO J. 5:2377-2385 (1986); Kleinerman et al., Cancer Res. 55:2831-2836 (1995); Hamada et al., Gynecol. Oncol., 63:219-227 (1996); Nabel et al., Science, 249:1285-1288 (1990); and Berkner,
10 BioTechniques 6:616-629 (1989), each of which is incorporated herein by reference).

In addition to a viral vector, a nucleic acid molecule of the invention can be administered using a liposome. Methods of making a liposome containing a
15 nucleic acid molecule are known in the art (see, Nabel et al., Proc. Natl. Acad. Sci. USA, 90:11307-11311 (1993), which is incorporated herein by reference; and Nabel et al., *supra*, 1990). Such liposomes can be made target specific by incorporating, for example, lipid-conjugated
20 antibodies into the structure of the liposome (see, Holmberg et al., J. Liposome Res., 1:393-406 (1990), which is incorporated herein by reference) or by incorporating a ligand or a receptor, that is bound by a corresponding receptor or ligand, respectively, that is
25 expressed on the target cell.

The present invention also provides methods of identifying the presence of an MG50 melanoma associated antigen in an individual. Such a method can be performed, for example, by contacting a biological sample
30 obtained from the individual with an antibody that specifically binds an MG50 epitope, wherein specific binding of the antibody to a component of the sample identifies the presence of the MG50 melanoma associated antigen in the individual. Such a biological sample can
35 be, for example, a tissue or tumor sample, which can be

obtained by a biopsy procedure from an individual suspected of having a cancer in which the cancer cells express an MG50 melanoma associated antigen.

In addition, the invention provides methods of
5 identifying the presence in an immune response against an MG50 melanoma associated antigen in an individual. Such a method can be performed, for example, by contacting a biological sample obtained from the subject with a peptide comprising at least six contiguous amino acids,
10 generally at least 8 contiguous amino acids, encoded by SEQ ID NO: 1 and detecting an immunoeffector function of the sample due to contact with the peptide, thereby identifying the presence of an immune response against an MG50 melanoma associated antigen in the individual. An
15 immunoeffector function can be, for example, the presence of anti-MG50 antibodies in the biological sample or the presence of MG50 reactive T cells in the sample. A biological sample can be, for example, a blood sample or a lymph tissue sample. For example, the peptide can be
20 an MG50 T cell epitope and the ability of epitope to stimulate the proliferation of T cells, which are a component of the biological sample, identifies the presence of MG50 reactive T cells in the sample and, therefore, the presence of an immune response against
25 MG50 in the individual from whom the sample was obtained.

The following examples are provided to illustrate embodiments of the invention.

EXAMPLE I

NUCLEIC ACID MOLECULE ENCODING MG50

30 This example describes methods for obtaining a nucleic acid molecule encoding the MG50 melanoma associated antigen.

Subtractive hybridization of cDNA obtained from melanoma cell line MSM M-1 ("M1") against an excess of mRNA from a squamous lung carcinoma cell line Lu-1 was used to clone cDNA sequences differentially expressed in the M1 melanoma cells (Hutchins et al., Cancer Research 51:1418-1425 (1991), which is incorporated herein by reference). Twelve candidate differentially expressed clones were obtained, six of which were considered novel based on a lack of sequence homology to sequence in the GenBank database (Hutchins et al., *supra*, 1991). One of these six clones, designated "melanoma gene-50" ("MG50") was selected for further characterization.

Based on northern blot analysis using the MG50 cDNA as a probe, MG50 is encoded by an mRNA of about 8.1 kilobases (kb). MG50 mRNA was detected in melanoma cells, lung carcinoma cells, rhabdomyosarcoma cells, fetal brain, fetal heart and human placenta.

Figure 1 shows a 6848 nucleotide portion of the cDNA encoding MG50 (SEQ ID NO: 1). Nucleotides 5510 to 6848 of SEQ ID NO: 1) were reported previously (Weiler, *supra*, 1993) However, efforts to continue sequencing the cDNA in the 5' direction largely were unsuccessful. As disclosed herein, primers were made based on the 5' end of the portion of the sequence described by Weiler (*supra*, 1993) and used to obtain more 5' sequences, which then were sequenced and used to search the Merck EST database. Overlapping EST sequences were identified and used to extend the sequence of MG50 to nucleotide 4336 of SEQ ID NO: 1 (i.e., nucleotides 4336 to 5509). Additional 5' sequences were determined by anchored PCR, RACE and DNA sequencing to nucleotide 3555 of SEQ ID NO: 1. Recently, the portion SEQ ID NO: 1 shown as nucleotides 1 to 5510 was submitted to GenBank as Accession No. D86983 and, therefore, nucleotides 1 to

3554 were added to produce the sequence shown as SEQ ID NO: 1 (Figures 1A to 1E).

Based on the 8.1 kb mRNA for MG50, it is estimated that approximately 1300 nucleotides remain to be sequenced to obtain the full length MG50 cDNA sequence. At least some of the 1300 nucleotides are expected to be 5' to the sequence shown in SEQ ID NO: 1, since an ATG initiation codon has not yet been identified. The remaining MG50 cDNA sequences can be obtained, for example, using a PCR method such as a RACE method.

A deduced amino acid sequence encoded by nucleotides 1 to 4488 of the MG50 cDNA is shown in Figure 2 (SEQ ID NO: 2). The 1496 amino acid polypeptide shares homology to Drosophila peroxidase and to products of the human peroxidase gene family. The 1496 amino acid MG50 polypeptide (SEQ ID NO: 2) is shown because the first stop codon encoded by SEQ ID NO: 1 occurs at nucleotides 4489 to 4491. However, if additional amino acids are deduced beyond this stop codon, cryptic coding sequences are revealed. Although additional stop codons are interspersed throughout the cryptic coding region, the sequence is considered to be a cryptic coding region because the peptide RPRPEQEPLP (SEQ ID NO: 4), which is encoded by nucleotides 5410 to 5439 of SEQ ID NO: 1 has the characteristics of an MG50 T cell epitope. Specifically, the peptide of SEQ ID NO: 4 binds more strongly to HLA-B7 than any other epitope tested in a competitive binding assay and stimulates proliferation of CD8⁺ T cells that were specifically reactive with melanoma cells expressing MG50 (see Example II).

Remarkably, the coding sequence of SEQ ID NO: 4 is downstream of nine stop codons, including the stop codon at nucleotides 4489 to 4491. While the presence of

this T cell epitope (SEQ ID NO: 4) in the cryptic region of SEQ ID NO: 1 may be fortuitous, another possibility is that the cryptic coding region, or a portion of this region, is translated in melanoma cells.

5

EXAMPLE II**POTENTIAL MG50 T CELL EPITOPES**

This example provides peptides that are encoded by SEQ ID NO: 1 and have characteristics of MHC class I restricted T cell epitopes.

10

The peptide RPRPEQEPLP (SEQ ID NO: 4), which is encoded by nucleotides 5410 to 5439 of SEQ ID NO: 1, in the cryptic region, stimulated proliferation of CD8⁺ T cells that were specifically reactive with melanoma cells expressing MG50 (see Example II). In these

15

experiments, T2 cells were transduced to express HLA-B7, then incubated with the peptide and CD8⁺ T cells (Tc cells), which were obtained from a patient having a melanoma that expressed MG50. Proliferation of the T cells indicated that the RPRPEQEPLP (SEQ ID NO: 4)

20

peptide acts as an HLA-B7 restricted Tc cell epitope.

In other experiments, CD8 T cells were generated against the RPRPEQEPLP (SEQ ID NO: 4) peptide *in vitro*, then reacted against RPRPEQEPLP (SEQ ID NO: 4) pulsed Cos-7 cells, which were transduced to express HLA-B7. The Cos-7 cells were lysed, demonstrating that the Tc cells recognized the MG50 T cell epitope in the context of HLA-B7 and were specifically reactive for the peptide.

25

Potential T cell epitopes encoded within the open reading frame of SEQ ID NO: 1 (i.e., present within SEQ ID NO: 2) were identified by homology to consensus HLA-A1 and HLA-A2 epitope sequences (see Kaat et al.,

30

supra, 1994; Falk and Rotzschko, *supra*, 1993). As shown in Table 1, a peptide having that characteristics expected of an HLA-A1 epitope (SEQ ID NO: 5) was identified and 12 peptides having characteristics of an HLA-A2 epitope (SEQ ID NOS: 6-17) were identified.

TABLE 1		
Amino Acid Sequence	*Amino Acid Position	SEQ ID NO:
CSEQPFPEHTASVQHAD	**	3
RPRPEQEPLP	1801-1810	4
DVTSGNTVY	273-281	5
VLFCAWGTL	34-42	6
CMHLLLEAV	66-74	7
LLLEAVPAV	69-77	8
TLHCDCEIL	210-218	9
VLSVNVPDV	625-633	10
DL DSTVVAL	845-853	11
WLPKILGEV	1051-1059	12
PLL RGLFGV	1133-1141	13
RLGPTLMCL	1244-1252	14
LLSTQFKRL	1252-1260	15
EMQKTITDL	1408-1416	16
DLRTQIKKL	1415-1423	17

* - amino acid position with respect to SEQ ID NO: 2.

** - not encoded by same reading frame as SEQ ID NO: 2.

The ability of the peptides shown in Table 1, or other potential T cell epitope encoded by SEQ ID NO: 1, to act as an MG50 T cell epitope can be determined using a T cell proliferation assay as described above. In addition, Cos-7 cells can be cotransduced with a cDNA

encoding HLA-A1 or HLA-A2, as appropriate, and with a nucleic acid molecule encoding a potential T cell epitope. The cotransduced Cos-7 cells then can be incubated with Tc cells that are specifically reactive
5 with melanoma cells expressing MG50 and MG50 T cell epitopes can be identified by detecting lysis of the Cos-7 cells.

Although the invention has been described with reference to the example provided above, it should be
10 understood that various changes can be made without departing from the spirit of the invention. Accordingly, the invention is limited only by the claims.

SEQUENCE LISTING

(1) GENERAL INFORMATION:

- (i) APPLICANT: The Regents of the University of California
The University of Southern California
- (ii) TITLE OF INVENTION: A Melanoma Associated Antigen, T Cell
Epitopes Thereof and Methods of Using Same
- (iii) NUMBER OF SEQUENCES: 26
- (iv) CORRESPONDENCE ADDRESS:
 - (A) ADDRESSEE: Campbell & Flores LLP
 - (B) STREET: 4370 La Jolla Village Drive, Suite 700
 - (C) CITY: San Diego
 - (D) STATE: California
 - (E) COUNTRY: United States
 - (F) ZIP: 92122
- (v) COMPUTER READABLE FORM:
 - (A) MEDIUM TYPE: Floppy disk
 - (B) COMPUTER: IBM PC compatible
 - (C) OPERATING SYSTEM: PC-DOS/MS-DOS
 - (D) SOFTWARE: PatentIn Release #1.0, Version #1.25
- (vi) CURRENT APPLICATION DATA:
 - (A) APPLICATION NUMBER:
 - (B) FILING DATE:
 - (C) CLASSIFICATION:
- (vii) PRIOR APPLICATION DATA:
 - (A) APPLICATION NUMBER: US 08/870,941
 - (B) FILING DATE: 06-JUN-1997
- (viii) ATTORNEY/AGENT INFORMATION:
 - (A) NAME: Campbell, Cathryn A.
 - (B) REGISTRATION NUMBER: 31,815
 - (C) REFERENCE/DOCKET NUMBER: FP-UD 3175
- (ix) TELECOMMUNICATION INFORMATION:
 - (A) TELEPHONE: (619) 535-9001
 - (B) TELEFAX: (619) 535-8949

(2) INFORMATION FOR SEQ ID NO:1:

- (i) SEQUENCE CHARACTERISTICS:
 - (A) LENGTH: 6847 base pairs
 - (B) TYPE: nucleic acid
 - (C) STRANDEDNESS: double
 - (D) TOPOLOGY: linear
- (ii) MOLECULE TYPE: DNA (genomic)
- (ix) FEATURE:
 - (A) NAME/KEY: CDS
 - (B) LOCATION: 1..4489

(xi) SEQUENCE DESCRIPTION: SEQ ID NO:1:

AGC Ser 1	CGG Arg	CCG Pro	TGG Trp	TGG Trp 5	CTC Leu	CGT Arg	GCG Ala	TCC Ser	GAG Glu 10	CGT Arg	CCG Pro	TCC Ser	GCG Ala	CCG Pro 15	TCG Ser	48
GCC Ala	ATG Met	GCC Ala	AAG Lys 20	CGC Arg	TCC Ser	AGG Arg	GGC Gly	CCC Pro 25	GGG Gly	CGC Arg	CGC Arg	TGC Cys	CTG Leu 30	TTG Leu	GCG Ala	96
CTC Leu	GTG Val	CTG Leu 35	TTC Phe	TGC Cys	GCC Ala	TGG Trp	GGG Gly 40	ACG Thr	CTG Leu	GCC Ala	GTG Val	GTG Val 45	GCC Ala	CAG Gln	AAG Lys	144
CCG Pro	GGC Gly 50	GCA Ala	GGG Gly	TGT Cys	CCG Pro	AGC Ser 55	CGC Arg	TGC Cys	CTG Leu	TGC Cys	TTC Phe 60	CGC Arg	ACC Thr	ACC Thr	GTG Val	192
CGC Arg 65	TGC Cys	ATG Met	CAT His	CTG Leu 70	CTG Leu	CTG Leu	GAG Glu	GCC Ala	GTG Val	CCC Pro 75	GCC Ala	GTG Val	GCG Ala	CCG Pro	CAG Gln 80	240
ACC Thr	TCC Ser	ATC Ile	CTA Leu	GAT Asp 85	CTT Leu	CGC Arg	TTT Phe	AAC Asn	AGA Arg 90	ATC Ile	AGA Arg	GAG Glu	ATC Ile	CAA Gln 95	CCT Pro	288
GGG Gly	GCA Ala	TTC Phe	AGG Arg 100	CGG Arg	CTG Leu	AGG Arg	AAC Asn	TTG Leu 105	AAC Asn	ACA Thr	TTG Leu	CTT Leu	CTC Leu 110	AAT Asn	AAT Asn	336
AAT Asn	CAG Gln 115	ATC Ile	AAG Lys	AGG Arg	ATA Ile	CCT Pro	AGT Ser 120	GGA Gly	GCA Ala	TTT Phe	GAA Glu 125	GAC Asp	TTG Leu	GAA Glu	AAT Asn	384
TTA Leu 130	AAA Lys	TAT Tyr	CTC Leu	TAT Tyr	CTG Leu	TAC Tyr 135	AAG Lys	AAT Asn	GAG Glu	ATC Ile	CAG Gln 140	TCA Ser	ATT Ile	GAC Asp	AGG Arg	432
CAA Gln 145	GCA Ala	TTT Phe	AAG Lys	GGA Gly	CTT Leu 150	GCC Ala	TCT Ser	CTA Leu	GAG Glu	CAA Gln 155	CTA Leu	TAC Tyr	CTG Leu	CAC His	TTT Phe 160	480
AAT Asn	CAG Gln	ATA Ile	GAA Glu	ACT Thr 165	TTG Leu	GAC Asp	CCA Pro	GAT Asp	TCG Ser 170	TTC Phe	CAG Gln	CAT His	CTC Leu	CCG Pro 175	AAG Lys	528
CTC Leu	GAG Glu	AGG Arg	CTA Leu 180	TTT Phe	TTG Leu	CAT His	AAC Asn 185	AAC Asn	CGG Arg	ATT Ile	ACA Thr	CAT His	TTA Leu 190	GTT Val	CCA Pro	576
GGG Gly	ACA Thr 195	TTT Phe	AAT Asn	CAC His	TTG Leu	GAA Glu	TCT Ser 200	ATG Met	AAG Lys	AGA Arg	TTG Leu	CGA Arg 205	CTG Leu	GAC Asp	TCA Ser	624
AAC Asn 210	ACA Thr	CTT Leu	CAC His	TGC Cys	GAC Asp	TGT Cys 215	GAA Glu	ATC Ile	CTG Leu	TGG Trp	TTG Leu 220	GCG Ala	GAT Asp	TTG Leu	CTG Leu	672
AAA Lys 225	ACC Thr	TAC Tyr	GCG Ala	GAG Glu	TCG Ser 230	GGG Gly	AAC Asn	GCG Ala	CAG Gln	GCA Ala 235	GCG Ala	GCC Ala	ATC Ile	TGT Cys	GAA Glu 240	720

TAT	CCC	AGA	CGC	ATC	CAG	GGA	CGC	TCA	GTG	GCA	ACC	ATC	ACC	CCG	GAA	768
Tyr	Pro	Arg	Arg	Ile	Gln	Gly	Arg	Ser	Val	Ala	Thr	Ile	Thr	Pro	Glu	
				245					250					255		
GAG	CTG	AAC	TGT	GAA	AGG	CCC	CGG	ATC	ACC	TCC	GAG	CCC	CAG	GAC	GCA	816
Glu	Leu	Asn	Cys	Glu	Arg	Pro	Arg	Ile	Thr	Ser	Glu	Pro	Gln	Asp	Ala	
			260					265					270			
GAT	GTG	ACC	TCG	GGG	AAC	ACC	GTG	TAC	TTC	ACC	TGC	AGA	GCC	GAA	GGC	864
Asp	Val	Thr	Ser	Gly	Asn	Thr	Val	Tyr	Phe	Thr	Cys	Arg	Ala	Glu	Gly	
		275					280					285				
AAC	CCC	AAG	CCT	GAG	ATC	ATC	TGG	CTG	CGA	AAC	AAT	AAT	GAG	CTG	AGC	912
Asn	Pro	Lys	Pro	Glu	Ile	Ile	Trp	Leu	Arg	Asn	Asn	Asn	Glu	Leu	Ser	
	290					295					300					
ATG	AAG	ACA	GAT	TCC	CGC	CTA	AAC	TTG	CTG	GAC	GAT	GGG	ACC	CTG	ATG	960
Met	Lys	Thr	Asp	Ser	Arg	Leu	Asn	Leu	Leu	Asp	Asp	Gly	Thr	Leu	Met	
305					310					315					320	
ATC	CAG	AAC	ACA	CAG	GAG	ACA	GAC	CAG	GGT	ATC	TAC	CAG	TGC	ATG	GCA	1008
Ile	Gln	Asn	Thr	Gln	Glu	Thr	Asp	Gln	Gly	Ile	Tyr	Gln	Cys	Met	Ala	
				325					330				335			
AAG	AAC	GTG	GCC	GGA	GAG	GTG	AAG	ACG	CAA	GAG	GTG	ACC	CTC	AGG	TAC	1056
Lys	Asn	Val	Ala	Gly	Glu	Val	Lys	Thr	Gln	Glu	Val	Thr	Leu	Arg	Tyr	
			340					345					350			
TTC	GGG	TCT	CCA	GCT	CGA	CCC	ACT	TTT	GTA	ATC	CAG	CCA	CAG	AAT	ACA	1104
Phe	Gly	Ser	Pro	Ala	Arg	Pro	Thr	Phe	Val	Ile	Gln	Pro	Gln	Asn	Thr	
		355					360					365				
GAG	GTG	CTG	GTT	GGG	GAG	AGC	GTC	ACG	CTG	GAG	TGC	AGC	GCC	ACA	GGC	1152
Glu	Val	Leu	Val	Gly	Glu	Ser	Val	Thr	Leu	Glu	Cys	Ser	Ala	Thr	Gly	
	370					375					380					
CAC	CCC	CCG	CCG	CGG	ATC	TCC	TGG	ACG	AGA	GGT	GAC	CGC	ACA	CCC	TTG	1200
His	Pro	Pro	Pro	Arg	Ile	Ser	Trp	Thr	Arg	Gly	Asp	Arg	Thr	Pro	Leu	
385					390					395					400	
CCA	GTT	GAC	CCG	CGG	GTG	AAC	ATC	ACG	CCT	TCT	GGC	GGG	CTT	TAC	ATA	1248
Pro	Val	Asp	Pro	Arg	Val	Asn	Ile	Thr	Pro	Ser	Gly	Gly	Leu	Tyr	Ile	
				405					410					415		
CAG	AAC	GTC	GTA	CAG	GGG	GAC	AGC	GGA	GAG	TAT	GCG	TGC	TCT	GCG	ACC	1296
Gln	Asn	Val	Val	Gln	Gly	Asp	Ser	Gly	Glu	Tyr	Ala	Cys	Ser	Ala	Thr	
			420					425					430			
AAC	AAC	ATT	GAC	AGC	GTC	CAT	GCC	ACC	GCT	TTC	ATC	ATC	GTC	CAG	GCT	1344
Asn	Asn	Ile	Asp	Ser	Val	His	Ala	Thr	Ala	Phe	Ile	Ile	Val	Gln	Ala	
		435					440					445				
CTT	CCT	CAG	TTC	ACT	GTG	ACG	CCT	CAG	GAC	AGA	GTC	GTT	ATT	GAG	GGC	1392
Leu	Pro	Gln	Phe	Thr	Val	Thr	Pro	Gln	Asp	Arg	Val	Val	Ile	Glu	Gly	
	450					455					460					
CAG	ACC	GTG	GAT	TTC	CAG	TGT	GAA	GCC	AAG	GGC	AAC	CCG	CCG	CCC	GTC	1440
Gln	Thr	Val	Asp	Phe	Gln	Cys	Glu	Ala	Lys	Gly	Asn	Pro	Pro	Pro	Val	
465					470					475					480	
ATC	GCC	TGG	ACC	AAG	GGA	GGG	AGC	CAG	CTC	TCC	GTG	GAC	CGG	CGG	CAC	1488
Ile	Ala	Trp	Thr	Lys	Gly	Gly	Ser	Gln	Leu	Ser	Val	Asp	Arg	Arg	His	
				485					490					495		

40

CTG	GTC	CTG	TCA	TCG	GGA	ACA	CTT	AGA	ATC	TCT	GGT	GTT	GCC	CTC	CAC	1536
Leu	Val	Leu	Ser	Ser	Gly	Thr	Leu	Arg	Ile	Ser	Gly	Val	Ala	Leu	His	
			500					505					510			
GAC	CAG	GGC	CAG	TAC	GAA	TGC	CAG	GCT	GTC	AAC	ATC	ATC	GGC	TCC	CAG	1584
Asp	Gln	Gly	Gln	Tyr	Glu	Cys	Gln	Ala	Val	Asn	Ile	Ile	Gly	Ser	Gln	
		515					520					525				
AAG	GTC	GTG	GCC	CAC	CTG	ACT	GTG	CAG	CCC	AGA	GTC	ACC	CCA	GTG	TTT	1632
Lys	Val	Val	Ala	His	Leu	Thr	Val	Gln	Pro	Arg	Val	Thr	Pro	Val	Phe	
	530					535					540					
GCC	AGC	ATT	CCC	AGC	GAC	ACA	ACA	GTG	GAG	GTG	GGC	GCC	AAT	GTG	CAG	1680
Ala	Ser	Ile	Pro	Ser	Asp	Thr	Thr	Val	Glu	Val	Gly	Ala	Asn	Val	Gln	
545					550					555					560	
CTC	CCG	TGC	AGC	TCC	CAG	GGC	GAG	CCC	GAG	CCA	GCC	ATC	ACC	TGG	AAC	1728
Leu	Pro	Cys	Ser	Ser	Gln	Gly	Glu	Pro	Glu	Pro	Ala	Ile	Thr	Trp	Asn	
				565					570					575		
AAG	GAT	GGG	GTT	CAG	GTG	ACA	GAA	AGT	GGA	AAA	TTT	CAC	ATC	AGC	CCT	1776
Lys	Asp	Gly	Val	Gln	Val	Thr	Glu	Ser	Gly	Lys	Phe	His	Ile	Ser	Pro	
			580					585					590			
GAA	GGA	TTC	TTG	ACC	ATC	AAT	GAC	GTT	GGC	CCT	GCA	GAC	GCA	GGT	CGC	1824
Glu	Gly	Phe	Leu	Thr	Ile	Asn	Asp	Val	Gly	Pro	Ala	Asp	Ala	Gly	Arg	
		595					600					605				
TAT	GAG	TGT	GTG	GCC	CGG	AAC	ACC	ATT	GGG	TCG	GCC	TCG	GTG	AGC	ATG	1872
Tyr	Glu	Cys	Val	Ala	Arg	Asn	Thr	Ile	Gly	Ser	Ala	Ser	Val	Ser	Met	
	610					615					620					
GTG	CTC	AGT	GTG	AAC	GTT	CCT	GAC	GTC	AGT	CGA	AAT	GGA	GAT	CCG	TTT	1920
Val	Leu	Ser	Val	Asn	Val	Pro	Asp	Val	Ser	Arg	Asn	Gly	Asp	Pro	Phe	
625					630					635					640	
GTA	GCT	ACC	TCC	ATC	GTG	GAA	GCG	ATT	GCG	ACT	GTT	GAC	AGA	GCT	ATA	1968
Val	Ala	Thr	Ser	Ile	Val	Glu	Ala	Ile	Ala	Thr	Val	Asp	Arg	Ala	Ile	
				645					650					655		
AAC	TCA	ACC	CGA	ACA	CAT	TTG	TTT	GAC	AGC	CGT	CCT	CGT	TCT	CCA	AAT	2016
Asn	Ser	Thr	Arg	Thr	His	Leu	Phe	Asp	Ser	Arg	Pro	Arg	Ser	Pro	Asn	
			660					665					670			
GAT	TTG	CTG	GCC	TTG	TTC	CGG	TAT	CCG	AGG	GAT	CCT	TAC	ACA	GTT	GAA	2064
Asp	Leu	Leu	Ala	Leu	Phe	Arg	Tyr	Pro	Arg	Asp	Pro	Tyr	Thr	Val	Glu	
		675					680					685				
CAG	GCA	CGG	GCG	GGA	GAA	ATC	TTT	GAA	CGG	ACA	TTG	CAG	CTC	ATT	CAG	2112
Gln	Ala	Arg	Ala	Gly	Glu	Ile	Phe	Glu	Arg	Thr	Leu	Gln	Leu	Ile	Gln	
	690					695					700					
GAG	CAT	GTA	CAG	CAT	GGC	TTG	ATG	GTC	GAC	CTC	AAC	GGA	ACA	AGT	TAC	2160
Glu	His	Val	Gln	His	Gly	Leu	Met	Val	Asp	Leu	Asn	Gly	Thr	Ser	Tyr	
705					710					715					720	
CAC	TAC	AAC	GAC	CTG	GTG	TCT	CCA	CAG	TAC	CTG	AAC	CTC	ATC	GCA	AAC	2208
His	Tyr	Asn	Asp	Leu	Val	Ser	Pro	Gln	Tyr	Leu	Asn	Leu	Ile	Ala	Asn	
				725					730					735		
CTG	TCG	GGC	TGT	ACC	GCC	CAC	CGG	CGC	GTG	AAC	AAC	TGC	TCG	GAC	ATG	2256
Leu	Ser	Gly	Cys	Thr	Ala	His	Arg	Arg	Val	Asn	Asn	Cys	Ser	Asp	Met	
			740					745						750		

TGC	TTC	CAC	CAG	AAG	TAC	CGG	ACG	CAC	GAC	GGC	ACC	TGT	AAC	AAC	CTG	2304
Cys	Phe	His	Gln	Lys	Tyr	Arg	Thr	His	Asp	Gly	Thr	Cys	Asn	Asn	Leu	
		755					760					765				
CAG	CAC	CCC	ATG	TGG	GGC	GCC	TCG	CTG	ACC	GCC	TTC	GAG	CGC	CTG	CTG	2352
Gln	His	Pro	Met	Trp	Gly	Ala	Ser	Leu	Thr	Ala	Phe	Glu	Arg	Leu	Leu	
		770				775					780					
AAA	TCC	GTG	TAC	GAG	AAT	GGC	TTC	AAC	ACC	CCT	CGG	GGC	ATC	AAC	CCC	2400
Lys	Ser	Val	Tyr	Glu	Asn	Gly	Phe	Asn	Thr	Pro	Arg	Gly	Ile	Asn	Pro	
785					790					795					800	
CAC	CGA	CTG	TAC	AAC	GGG	CAC	GCC	CTT	CCC	ATG	CCG	CGC	CTG	GTG	TCC	2448
His	Arg	Leu	Tyr	Asn	Gly	His	Ala	Leu	Pro	Met	Pro	Arg	Leu	Val	Ser	
				805					810					815		
ACC	ACC	CTG	ATC	GGG	ACG	GAG	ACC	GTC	ACA	CCC	GAC	GAG	CAG	TTC	ACC	2496
Thr	Thr	Leu	Ile	Gly	Thr	Glu	Thr	Val	Thr	Pro	Asp	Glu	Gln	Phe	Thr	
			820					825						830		
CAC	ATG	CTG	ATG	CAG	TGG	GGC	CAG	TTC	CTG	GAC	CAC	GAC	CTC	GAC	TCC	2544
His	Met	Leu	Met	Gln	Trp	Gly	Gln	Phe	Leu	Asp	His	Asp	Leu	Asp	Ser	
		835					840					845				
ACG	GTG	GTG	GCC	CTG	AGC	CAG	GCA	CGC	TTC	TCC	GAC	GGA	CAG	CAC	TGC	2592
Thr	Val	Val	Ala	Leu	Ser	Gln	Ala	Arg	Phe	Ser	Asp	Gly	Gln	His	Cys	
		850				855					860					
AGC	AAC	GTG	TGC	AGC	AAC	GAC	CCC	CCC	TGC	TTC	TCT	GTC	ATG	ATC	CCC	2640
Ser	Asn	Val	Cys	Ser	Asn	Asp	Pro	Pro	Cys	Phe	Ser	Val	Met	Ile	Pro	
865					870					875					880	
CCC	AAT	GAC	TCC	CGG	GCC	AGG	AGC	GGG	GCC	CGC	TGC	ATG	TTC	TTC	GTG	2688
Pro	Asn	Asp	Ser	Arg	Ala	Arg	Ser	Gly	Ala	Arg	Cys	Met	Phe	Phe	Val	
				885					890					895		
CGC	TCC	AGC	CCT	GTG	TGC	GGC	AGC	GGC	ATG	ACT	TCG	CTG	CTC	ATG	AAC	2736
Arg	Ser	Ser	Pro	Val	Cys	Gly	Ser	Gly	Met	Thr	Ser	Leu	Leu	Met	Asn	
			900					905					910			
TCC	GTG	TAC	CCG	CGG	GAG	CAG	ATC	AAC	CAG	CTC	ACC	TCC	TAC	ATC	GAC	2784
Ser	Val	Tyr	Pro	Arg	Glu	Gln	Ile	Asn	Gln	Leu	Thr	Ser	Tyr	Ile	Asp	
		915					920					925				
GCA	TCC	AAC	GTG	TAC	GGG	AGC	ACG	GAG	CAT	GAG	GCC	CGC	AGC	ATC	CGC	2832
Ala	Ser	Asn	Val	Tyr	Gly	Ser	Thr	Glu	His	Glu	Ala	Arg	Ser	Ile	Arg	
		930				935					940					
GAC	CTG	GCC	AGC	CAC	CGC	GGC	CTG	CTG	CGG	CAG	GGC	ATC	GTG	CAG	CGG	2880
Asp	Leu	Ala	Ser	His	Arg	Gly	Leu	Leu	Arg	Gln	Gly	Ile	Val	Gln	Arg	
945					950					955					960	
TCC	GGG	AAG	CCG	CTG	CTC	CCC	TTC	GCC	ACC	GGG	CCG	CCC	ACG	GAG	TGC	2928
Ser	Gly	Lys	Pro	Leu	Leu	Pro	Phe	Ala	Thr	Gly	Pro	Pro	Thr	Glu	Cys	
				965					970					975		
ATG	CGG	GAC	GAG	AAC	GAG	AGC	CCC	ATC	CCC	TGC	TTC	CTG	GCC	GGG	GAC	2976
Met	Arg	Asp	Glu	Asn	Glu	Ser	Pro	Ile	Pro	Cys	Phe	Leu	Ala	Gly	Asp	
			980					985					990			
CAC	CGC	GCC	AAC	GAG	CAG	CTG	GGC	CTG	ACC	AGC	ATG	CAC	ACG	CTG	TGG	3024
His	Arg	Ala	Asn	Glu	Gln	Leu	Gly	Leu	Thr	Ser	Met	His	Thr	Leu	Trp	
		995					1000						1005			

42

TTC CGC GAG CAC AAC CGC ATT GCC ACG GAG CTG CTC AAG CTG AAC CCG Phe Arg Glu His Asn Arg Ile Ala Thr Glu Leu Leu Lys Leu Asn Pro 1010 1015 1020	3072
CAC TGG GAC GGC GAC ACC ATC TAC TAT GAG ACC AGG AAG ATC GTG GGT His Trp Asp Gly Asp Thr Ile Tyr Tyr Glu Thr Arg Lys Ile Val Gly 1025 1030 1035 1040	3120
GCG GAG ATC CAG CAC ATC ACC TAC CAG CAC TGG CTC CCG AAG ATC CTG Ala Glu Ile Gln His Ile Thr Tyr Gln His Trp Leu Pro Lys Ile Leu 1045 1050 1055	3168
GGG GAG GTG GGC ATG AGG ACG CTG GGA GAG TAC CAC GGC TAC GAC CCC Gly Glu Val Gly Met Arg Thr Leu Gly Glu Tyr His Gly Tyr Asp Pro 1060 1065 1070	3216
GGC ATC AAT GCT GGC ATC TTC AAC GCC TTC GCC ACC GCG GCC TTC AGG Gly Ile Asn Ala Gly Ile Phe Asn Ala Phe Ala Thr Ala Ala Phe Arg 1075 1080 1085	3264
TTT GGC CAC ACG CTT GTC AAC CCA CTG CTT TAC CGG CTG GAC GAG AAC Phe Gly His Thr Leu Val Asn Pro Leu Leu Tyr Arg Leu Asp Glu Asn 1090 1095 1100	3312
TTC CAG CCC ATT GCA CAA GAT CAC CTC CCC CTT CAC AAA GCT TTC TTC Phe Gln Pro Ile Ala Gln Asp His Leu Pro Leu His Lys Ala Phe Phe 1105 1110 1115 1120	3360
TCT CCC TTC CGG ATT GTG AAT GAG GGC GGC ATC GAT CCG CTT CTC AGG Ser Pro Phe Arg Ile Val Asn Glu Gly Gly Ile Asp Pro Leu Leu Arg 1125 1130 1135	3408
GGG CTG TTC GGG GTG GCG GGG AAA ATG CGT GTG CCC TCG CAG CTG CTG Gly Leu Phe Gly Val Ala Gly Lys Met Arg Val Pro Ser Gln Leu Leu 1140 1145 1150	3456
AAC ACG GAG CTC ACG GAG CGG CTG TTC TCC ATG GCA CAC ACG GTG GCT Asn Thr Glu Leu Thr Glu Arg Leu Phe Ser Met Ala His Thr Val Ala 1155 1160 1165	3504
CTG GAC CTG GCG GCC ATC AAC ATC CAG CGG GGC CGG GAC CAC GGG ATC Leu Asp Leu Ala Ala Ile Asn Ile Gln Arg Gly Arg Asp His Gly Ile 1170 1175 1180	3552
CCA CCC TAC CAC GAC TAC AGG GTC TAC TGC AAT CTA TCG GCG GCA CAC Pro Pro Tyr His Asp Tyr Arg Val Tyr Cys Asn Leu Ser Ala Ala His 1185 1190 1195 1200	3600
ACG TTC GAG GAC CTG AAA AAT GAG ATT AAA AAC CCT GAG ATC CGG GAG Thr Phe Glu Asp Leu Lys Asn Glu Ile Lys Asn Pro Glu Ile Arg Glu 1205 1210 1215	3648
AAA CTG AAA AGG TTG TAT GGC TCG ACA CTC AAC ATC GAC CTG TTT CCG Lys Leu Lys Arg Leu Tyr Gly Ser Thr Leu Asn Ile Asp Leu Phe Pro 1220 1225 1230	3696
GCG CTC GTG GTG GAG GAC CTG GTG CCT GGC AGC CGG CTG GGC CCC ACC Ala Leu Val Val Glu Asp Leu Val Pro Gly Ser Arg Leu Gly Pro Thr 1235 1240 1245	3744
CTG ATG TGT CTT CTC AGC ACA CAG TTC AAG CGC CTG CGA GAT GGG GAC Leu Met Cys Leu Leu Ser Thr Gln Phe Lys Arg Leu Arg Asp Gly Asp 1250 1255 1260	3792

43

AGG TTG TGG TAT GAG AAC CCT GGG GTG TTC TCC CCG GCC CAG CTG ACT	3840
Arg Leu Trp Tyr Glu Asn Pro Gly Val Phe Ser Pro Ala Gln Leu Thr	
1265 1270 1275 1280	
CAG ATC AAG CAG ACG TCG CTG GCC AGG ATC CTA TGC GAC AAC GCG GAC	3888
Gln Ile Lys Gln Thr Ser Leu Ala Arg Ile Leu Cys Asp Asn Ala Asp	
1285 1290 1295	
AAC ATC ACC CGG GTG CAG AGC GAC GTG TTC AGG GTG GCG GAG TTC CCT	3936
Asn Ile Thr Arg Val Gln Ser Asp Val Phe Arg Val Ala Glu Phe Pro	
1300 1305 1310	
CAC GGC TAC GGC AGC TGT GAC GAG ATC CCC AGG GTG GAC CTC CGG GTG	3984
His Gly Tyr Gly Ser Cys Asp Glu Ile Pro Arg Val Asp Leu Arg Val	
1315 1320 1325	
TGG CAG GAC TGC TGT GAA GAC TGT AGG ACC AGG GGG CAG TTC AAT GCC	4032
Trp Gln Asp Cys Cys Glu Asp Cys Arg Thr Arg Gly Gln Phe Asn Ala	
1330 1335 1340	
TTT TCC TAT CAT TTC CGA GGC AGA CGG TCT CTT GAG TTC AGC TAC CAG	4080
Phe Ser Tyr His Phe Arg Gly Arg Arg Ser Leu Glu Phe Ser Tyr Gln	
1345 1350 1355 1360	
GAG GAC AAG CCG ACC AAG AAA ACA AGA CCA CGG AAA ATA CCC AGT GTT	4128
Glu Asp Lys Pro Thr Lys Lys Thr Arg Pro Arg Lys Ile Pro Ser Val	
1365 1370 1375	
GGG AGA CAG GGG GAA CAT CTC AGC AAC AGC ACC TCA GCC TTC AGC ACA	4176
Gly Arg Gln Gly Glu His Leu Ser Asn Ser Thr Ser Ala Phe Ser Thr	
1380 1385 1390	
CGC TCA GAT GCA TCT GGG ACA AAT GAC TTC AGA GAG TTT GTT CTG GAA	4224
Arg Ser Asp Ala Ser Gly Thr Asn Asp Phe Arg Glu Phe Val Leu Glu	
1395 1400 1405	
ATG CAG AAG ACC ATC ACA GAC CTC AGA ACA CAG ATA AAG AAA CTT GAA	4272
Met Gln Lys Thr Ile Thr Asp Leu Arg Thr Gln Ile Lys Lys Leu Glu	
1410 1415 1420	
TCA CGG CTC AGT ACC ACA GAG TGC GTG GAT GCC GGG GGC GAA TCT CAC	4320
Ser Arg Leu Ser Thr Thr Glu Cys Val Asp Ala Gly Gly Glu Ser His	
1425 1430 1435 1440	
GCC AAC AAC ACC AAG TGG AAA AAA GAT GCA TGC ACC ATT TGT GAA TGC	4368
Ala Asn Asn Thr Lys Trp Lys Lys Asp Ala Cys Thr Ile Cys Glu Cys	
1445 1450 1455	
AAA GAC GGG CAG GTC ACC TGC TTC GTG GAA GCT TGC CCC CCT GCC ACC	4416
Lys Asp Gly Gln Val Thr Cys Phe Val Glu Ala Cys Pro Pro Ala Thr	
1460 1465 1470	
TGT GCT GTC CCC GTG AAC ATC CCA GGG GCC TGC TGT CCA GTC TGC TTA	4464
Cys Ala Val Pro Val Asn Ile Pro Gly Ala Cys Cys Pro Val Cys Leu	
1475 1480 1485	
CAG AAG AGG GCG GAG GAA AAG CCC T AGGCTCCTGG GAGGCTCCTC	4509
Gln Lys Arg Ala Glu Glu Lys Pro	
1490 1495	
AGAGTTTGTC TGCTGTGCCA TCGTGAGATC GGGTGGCCGA TGGCAGGGAG CTGCGGACTG	4569
CAGACCAGGA AACACCCAGA ACTCGTGACA TTTCATGACA ACGTCCAGCT GGTGCTGTTA	4629

CAGAAGGCAG	TGCAGGAGGC	TTCCAACCAG	AGCATCTGCG	GAGAAGGAGG	CACAGCAGGT	4689
GCCTGAAGGG	AAGCAGGCAG	GAGTCCTAGC	TTCACGTTAG	ACTTCTCAGG	TTTTTATTTA	4749
ATTCTTTTAA	AATGAAAAAT	TGGTGCTACT	ATTAAATTGC	ACAGTTGAAT	CATTTAGGCG	4809
CCTAAATTGG	TTTTGCCTCC	CAACACCATT	TCTTTTTTAA	TAAAGCAGGA	TACCTCTATA	4869
TGTCAGCCTT	GCCTTGTTCA	GATGCCAGGA	GCCGGCAGAC	CTGTCACCCG	CAGGTGGGGT	4929
GAGTCTCGGA	GCTGCCAGAG	GGGCTCACCG	AAATCGGGGT	TCCATCACAA	GCTATGTTTA	4989
AAAAGAAAAT	TGGTGTTTGG	CAAACGGAAC	AGAACCTTTG	ATGAGAGCGT	TCACAGGGAC	5049
ACTGTCTGGG	GGTGCAGTGC	AAGCCCCCGG	CCTCTTCCCT	GGGAACCTCT	GAACCTCTCC	5109
TTCCTCTGGG	CTCTCTGTAA	CATTTACCA	CACGTCAGCA	TCTAATCCCA	AGACAAACAT	5169
TCCCGCTGCT	CGAAGCAGCT	GTATAGCCTG	TGACTCTCCG	TGTGTCAGCT	CCTTCCACAC	5229
CTGATTAGAA	CATTCATAAG	CCACATTTAG	AAACAGATTT	GCTTTCAGCT	GTCACCTGCA	5289
CACATACTGC	CTAGTTGTGA	ACCAAATGTG	AAAAAACCTC	CTTCATCCCA	TTGTGTATCT	5349
GATACCTGCC	GAGGGCCAAG	GGTGTGTGTT	GACAACGCCG	CTCCCAGCCG	GCCCTGGTTG	5409
CGTCCACGTC	CTGAACAAGA	GCCGCTTCCG	GATGGCTCTT	CCCAAGGGAG	GAGGAGCTCA	5469
AGTGTCTGGG	ACTGTCTAAC	TTCAGGTTGT	GTGAGTGCGT	TAAAAAAAAA	AAAAAAAAAA	5529
AGAATCCCTA	TACCTCATTT	GTATTTTTTA	AATGCGTGAT	GTTTTATGAA	ATTGTGTCCA	5589
TTTTTTAGGT	ATTAGATATG	GCAGAAAAAC	CATTTCCACT	ATGCAAAGTT	CTTTTAGACG	5649
TCAGTGAAAA	TCAACTCTCA	TACCTCATGG	GTCTCTCTTT	AATTGACCAA	AACCTTCCAT	5709
TTTTCTCTTA	AATACAAAGC	GATCTGTGTT	CTGAGCAACC	TTCCCCGAA	CACACAGCTT	5769
CAGTGCAGCA	CGCTGACCTG	AGTATCCACC	AGGTGCCAGG	CACAGTTGCT	GGGCNNACGG	5829
AGGCACCAAG	GTCCGGGCCA	CCTGCCCGCA	GGCAAGGCCC	AGCTGAGGTG	GTGGGAGGGG	5889
AGCCCCTGAG	GTCAGGGGCC	GTTTCGGTTC	AGGGTGGCAG	GTGTCCAGCA	CTGGGGTATG	5949
GCGTCGAGGC	TTCCATGGGG	TGGGGGAGGC	CAGCTTCCTT	CTGACAGGAT	GGGCGCATAC	6009
AGTGCCTGGT	GTGATTTGTG	CACAACCCGT	GTTCCAGGTG	CACATCCTCC	CAAGGAGACA	6069
CCCAGACCTT	TCCAGCACGG	GCCGGCCAAG	TTGCTGCGGC	GGAGGCAGCA	TTTCAGCTGT	6129
GAGGAAGGTC	ATTGGATTCA	TGTGTTTTAT	CTGTAAAAAT	GGTTGTCTTA	ACTTCTTAAC	6189
TCATATTGGT	AAGTGATTGA	TAAAAATTGG	TTGGTGTTTT	CATGACATGT	GGACTTCTNT	6249
TGNATAGAAG	TCAAATGTAG	TGACAATTTG	TGGAAGAGAT	TCTTGTCAAA	GTGAAATAGG	6309
AAATGTGTAA	GTTCGTCTAA	AAGCTGATGG	TTATGTAAGT	TGCTCAGGCA	CTCAGATGAC	6369
AGCAGATTCT	GGGTTCTGGG	AGTGTTCTGT	GCCTCTTACA	TGCCCTGGAG	GCCTCATGGT	6429
CTCAGTGCTG	AGGCGGCACA	CCTGTAGCAC	ACCTGCGTAA	TGTGCGGTCT	GGGCCAGTCA	6489
CAAGGAATTG	TGTTGTCTAA	NCCAAAGGGG	GAAGCTGACT	GTGTATTACC	AAAAAAAATT	6549

45

CTGTAATNCA AACCNAAATG TCTGCGGAAT CACCAGTTTG ATACTCTCTG TAATCAGAGC 6609
 AGTNGNCTGA GGGCGGNCAG TNCCTGGGTG AACGTGTCTA GCAGCCACTG TGGGGGATCG 6669
 CTGTAACAGG AGTGGAATGT ACATATTTAT TTACTTTTCT AACTGCTCCA ACAGCCAAAT 6729
 GCCTTTTTTTA TGACCATTTGT ATTCAGTTCA TTACCAAAGA AATGTTTGCA CTTTGTAATG 6789
 ATGCCTTTTCA GTTCAAATAA ATGGGTCACA TTTTCAAATG GAAAAAAAAA AAAAAAAAAA 6847

(2) INFORMATION FOR SEQ ID NO:2:

(i) SEQUENCE CHARACTERISTICS:

- (A) LENGTH: 1496 amino acids
- (B) TYPE: amino acid
- (D) TOPOLOGY: linear

(ii) MOLECULE TYPE: protein

(xi) SEQUENCE DESCRIPTION: SEQ ID NO:2:

Ser Arg Pro Trp Trp Leu Arg Ala Ser Glu Arg Pro Ser Ala Pro Ser
 1 5 10 15
 Ala Met Ala Lys Arg Ser Arg Gly Pro Gly Arg Arg Cys Leu Leu Ala
 20 25 30
 Leu Val Leu Phe Cys Ala Trp Gly Thr Leu Ala Val Val Ala Gln Lys
 35 40 45
 Pro Gly Ala Gly Cys Pro Ser Arg Cys Leu Cys Phe Arg Thr Thr Val
 50 55 60
 Arg Cys Met His Leu Leu Leu Glu Ala Val Pro Ala Val Ala Pro Gln
 65 70 75 80
 Thr Ser Ile Leu Asp Leu Arg Phe Asn Arg Ile Arg Glu Ile Gln Pro
 85 90 95
 Gly Ala Phe Arg Arg Leu Arg Asn Leu Asn Thr Leu Leu Leu Asn Asn
 100 105 110
 Asn Gln Ile Lys Arg Ile Pro Ser Gly Ala Phe Glu Asp Leu Glu Asn
 115 120 125
 Leu Lys Tyr Leu Tyr Leu Tyr Lys Asn Glu Ile Gln Ser Ile Asp Arg
 130 135 140
 Gln Ala Phe Lys Gly Leu Ala Ser Leu Glu Gln Leu Tyr Leu His Phe
 145 150 155 160
 Asn Gln Ile Glu Thr Leu Asp Pro Asp Ser Phe Gln His Leu Pro Lys
 165 170 175
 Leu Glu Arg Leu Phe Leu His Asn Asn Arg Ile Thr His Leu Val Pro
 180 185 190
 Gly Thr Phe Asn His Leu Glu Ser Met Lys Arg Leu Arg Leu Asp Ser
 195 200 205
 Asn Thr Leu His Cys Asp Cys Glu Ile Leu Trp Leu Ala Asp Leu Leu
 210 215 220

46

Lys Thr Tyr Ala Glu Ser Gly Asn Ala Gln Ala Ala Ala Ile Cys Glu
 225 230 235 240
 Tyr Pro Arg Arg Ile Gln Gly Arg Ser Val Ala Thr Ile Thr Pro Glu
 245 250 255
 Glu Leu Asn Cys Glu Arg Pro Arg Ile Thr Ser Glu Pro Gln Asp Ala
 260 265 270
 Asp Val Thr Ser Gly Asn Thr Val Tyr Phe Thr Cys Arg Ala Glu Gly
 275 280 285
 Asn Pro Lys Pro Glu Ile Ile Trp Leu Arg Asn Asn Asn Glu Leu Ser
 290 295 300
 Met Lys Thr Asp Ser Arg Leu Asn Leu Leu Asp Asp Gly Thr Leu Met
 305 310 315 320
 Ile Gln Asn Thr Gln Glu Thr Asp Gln Gly Ile Tyr Gln Cys Met Ala
 325 330 335
 Lys Asn Val Ala Gly Glu Val Lys Thr Gln Glu Val Thr Leu Arg Tyr
 340 345 350
 Phe Gly Ser Pro Ala Arg Pro Thr Phe Val Ile Gln Pro Gln Asn Thr
 355 360 365
 Glu Val Leu Val Gly Glu Ser Val Thr Leu Glu Cys Ser Ala Thr Gly
 370 375 380
 His Pro Pro Pro Arg Ile Ser Trp Thr Arg Gly Asp Arg Thr Pro Leu
 385 390 395 400
 Pro Val Asp Pro Arg Val Asn Ile Thr Pro Ser Gly Gly Leu Tyr Ile
 405 410 415
 Gln Asn Val Val Gln Gly Asp Ser Gly Glu Tyr Ala Cys Ser Ala Thr
 420 425 430
 Asn Asn Ile Asp Ser Val His Ala Thr Ala Phe Ile Ile Val Gln Ala
 435 440 445
 Leu Pro Gln Phe Thr Val Thr Pro Gln Asp Arg Val Val Ile Glu Gly
 450 455 460
 Gln Thr Val Asp Phe Gln Cys Glu Ala Lys Gly Asn Pro Pro Pro Val
 465 470 475 480
 Ile Ala Trp Thr Lys Gly Gly Ser Gln Leu Ser Val Asp Arg Arg His
 485 490 495
 Leu Val Leu Ser Ser Gly Thr Leu Arg Ile Ser Gly Val Ala Leu His
 500 505 510
 Asp Gln Gly Gln Tyr Glu Cys Gln Ala Val Asn Ile Ile Gly Ser Gln
 515 520 525
 Lys Val Val Ala His Leu Thr Val Gln Pro Arg Val Thr Pro Val Phe
 530 535 540
 Ala Ser Ile Pro Ser Asp Thr Thr Val Glu Val Gly Ala Asn Val Gln
 545 550 555 560

47

Leu Pro Cys Ser Ser Gln Gly Glu Pro Glu Pro Ala Ile Thr Trp Asn
 565 570 575
 Lys Asp Gly Val Gln Val Thr Glu Ser Gly Lys Phe His Ile Ser Pro
 580 585 590
 Glu Gly Phe Leu Thr Ile Asn Asp Val Gly Pro Ala Asp Ala Gly Arg
 595 600 605
 Tyr Glu Cys Val Ala Arg Asn Thr Ile Gly Ser Ala Ser Val Ser Met
 610 615 620
 Val Leu Ser Val Asn Val Pro Asp Val Ser Arg Asn Gly Asp Pro Phe
 625 630 635 640
 Val Ala Thr Ser Ile Val Glu Ala Ile Ala Thr Val Asp Arg Ala Ile
 645 650 655
 Asn Ser Thr Arg Thr His Leu Phe Asp Ser Arg Pro Arg Ser Pro Asn
 660 665 670
 Asp Leu Leu Ala Leu Phe Arg Tyr Pro Arg Asp Pro Tyr Thr Val Glu
 675 680 685
 Gln Ala Arg Ala Gly Glu Ile Phe Glu Arg Thr Leu Gln Leu Ile Gln
 690 695 700
 Glu His Val Gln His Gly Leu Met Val Asp Leu Asn Gly Thr Ser Tyr
 705 710 715 720
 His Tyr Asn Asp Leu Val Ser Pro Gln Tyr Leu Asn Leu Ile Ala Asn
 725 730 735
 Leu Ser Gly Cys Thr Ala His Arg Arg Val Asn Asn Cys Ser Asp Met
 740 745 750
 Cys Phe His Gln Lys Tyr Arg Thr His Asp Gly Thr Cys Asn Asn Leu
 755 760 765
 Gln His Pro Met Trp Gly Ala Ser Leu Thr Ala Phe Glu Arg Leu Leu
 770 775 780
 Lys Ser Val Tyr Glu Asn Gly Phe Asn Thr Pro Arg Gly Ile Asn Pro
 785 790 795 800
 His Arg Leu Tyr Asn Gly His Ala Leu Pro Met Pro Arg Leu Val Ser
 805 810 815
 Thr Thr Leu Ile Gly Thr Glu Thr Val Thr Pro Asp Glu Gln Phe Thr
 820 825 830
 His Met Leu Met Gln Trp Gly Gln Phe Leu Asp His Asp Leu Asp Ser
 835 840 845
 Thr Val Val Ala Leu Ser Gln Ala Arg Phe Ser Asp Gly Gln His Cys
 850 855 860
 Ser Asn Val Cys Ser Asn Asp Pro Pro Cys Phe Ser Val Met Ile Pro
 865 870 875 880
 Pro Asn Asp Ser Arg Ala Arg Ser Gly Ala Arg Cys Met Phe Phe Val
 885 890 895

48

Arg Ser Ser Pro Val Cys Gly Ser Gly Met Thr Ser Leu Leu Met Asn
 900 905 910
 Ser Val Tyr Pro Arg Glu Gln Ile Asn Gln Leu Thr Ser Tyr Ile Asp
 915 920 925
 Ala Ser Asn Val Tyr Gly Ser Thr Glu His Glu Ala Arg Ser Ile Arg
 930 935 940
 Asp Leu Ala Ser His Arg Gly Leu Leu Arg Gln Gly Ile Val Gln Arg
 945 950 955 960
 Ser Gly Lys Pro Leu Leu Pro Phe Ala Thr Gly Pro Pro Thr Glu Cys
 965 970 975
 Met Arg Asp Glu Asn Glu Ser Pro Ile Pro Cys Phe Leu Ala Gly Asp
 980 985 990
 His Arg Ala Asn Glu Gln Leu Gly Leu Thr Ser Met His Thr Leu Trp
 995 1000 1005
 Phe Arg Glu His Asn Arg Ile Ala Thr Glu Leu Leu Lys Leu Asn Pro
 1010 1015 1020
 His Trp Asp Gly Asp Thr Ile Tyr Tyr Glu Thr Arg Lys Ile Val Gly
 1025 1030 1035 1040
 Ala Glu Ile Gln His Ile Thr Tyr Gln His Trp Leu Pro Lys Ile Leu
 1045 1050 1055
 Gly Glu Val Gly Met Arg Thr Leu Gly Glu Tyr His Gly Tyr Asp Pro
 1060 1065 1070
 Gly Ile Asn Ala Gly Ile Phe Asn Ala Phe Ala Thr Ala Ala Phe Arg
 1075 1080 1085
 Phe Gly His Thr Leu Val Asn Pro Leu Leu Tyr Arg Leu Asp Glu Asn
 1090 1095 1100
 Phe Gln Pro Ile Ala Gln Asp His Leu Pro Leu His Lys Ala Phe Phe
 1105 1110 1115 1120
 Ser Pro Phe Arg Ile Val Asn Glu Gly Gly Ile Asp Pro Leu Leu Arg
 1125 1130 1135
 Gly Leu Phe Gly Val Ala Gly Lys Met Arg Val Pro Ser Gln Leu Leu
 1140 1145 1150
 Asn Thr Glu Leu Thr Glu Arg Leu Phe Ser Met Ala His Thr Val Ala
 1155 1160 1165
 Leu Asp Leu Ala Ala Ile Asn Ile Gln Arg Gly Arg Asp His Gly Ile
 1170 1175 1180
 Pro Pro Tyr His Asp Tyr Arg Val Tyr Cys Asn Leu Ser Ala Ala His
 1185 1190 1195 1200
 Thr Phe Glu Asp Leu Lys Asn Glu Ile Lys Asn Pro Glu Ile Arg Glu
 1205 1210 1215
 Lys Leu Lys Arg Leu Tyr Gly Ser Thr Leu Asn Ile Asp Leu Phe Pro
 1220 1225 1230

49

Ala Leu Val Val Glu Asp Leu Val Pro Gly Ser Arg Leu Gly Pro Thr
 1235 1240 1245

Leu Met Cys Leu Leu Ser Thr Gln Phe Lys Arg Leu Arg Asp Gly Asp
 1250 1255 1260

Arg Leu Trp Tyr Glu Asn Pro Gly Val Phe Ser Pro Ala Gln Leu Thr
 1265 1270 1275 1280

Gln Ile Lys Gln Thr Ser Leu Ala Arg Ile Leu Cys Asp Asn Ala Asp
 1285 1290 1295

Asn Ile Thr Arg Val Gln Ser Asp Val Phe Arg Val Ala Glu Phe Pro
 1300 1305 1310

His Gly Tyr Gly Ser Cys Asp Glu Ile Pro Arg Val Asp Leu Arg Val
 1315 1320 1325

Trp Gln Asp Cys Cys Glu Asp Cys Arg Thr Arg Gly Gln Phe Asn Ala
 1330 1335 1340

Phe Ser Tyr His Phe Arg Gly Arg Arg Ser Leu Glu Phe Ser Tyr Gln
 1345 1350 1355 1360

Glu Asp Lys Pro Thr Lys Lys Thr Arg Pro Arg Lys Ile Pro Ser Val
 1365 1370 1375

Gly Arg Gln Gly Glu His Leu Ser Asn Ser Thr Ser Ala Phe Ser Thr
 1380 1385 1390

Arg Ser Asp Ala Ser Gly Thr Asn Asp Phe Arg Glu Phe Val Leu Glu
 1395 1400 1405

Met Gln Lys Thr Ile Thr Asp Leu Arg Thr Gln Ile Lys Lys Leu Glu
 1410 1415 1420

Ser Arg Leu Ser Thr Thr Glu Cys Val Asp Ala Gly Gly Glu Ser His
 1425 1430 1435 1440

Ala Asn Asn Thr Lys Trp Lys Lys Asp Ala Cys Thr Ile Cys Glu Cys
 1445 1450 1455

Lys Asp Gly Gln Val Thr Cys Phe Val Glu Ala Cys Pro Pro Ala Thr
 1460 1465 1470

Cys Ala Val Pro Val Asn Ile Pro Gly Ala Cys Cys Pro Val Cys Leu
 1475 1480 1485

Gln Lys Arg Ala Glu Glu Lys Pro
 1490 1495

(2) INFORMATION FOR SEQ ID NO:3:

(i) SEQUENCE CHARACTERISTICS:

- (A) LENGTH: 17 amino acids
- (B) TYPE: amino acid
- (D) TOPOLOGY: linear

(ii) MOLECULE TYPE: peptide

(xi) SEQUENCE DESCRIPTION: SEQ ID NO:3:

Cys Ser Glu Gln Pro Phe Pro Glu His Thr Ala Ser Val Gln His Ala
1 5 10 15

Asp

(2) INFORMATION FOR SEQ ID NO:4:

(i) SEQUENCE CHARACTERISTICS:
(A) LENGTH: 10 amino acids
(B) TYPE: amino acid
(D) TOPOLOGY: linear

(ii) MOLECULE TYPE: peptide

(xi) SEQUENCE DESCRIPTION: SEQ ID NO:4:

Arg Pro Arg Pro Glu Gln Glu Pro Leu Pro
1 5 10

(2) INFORMATION FOR SEQ ID NO:5:

(i) SEQUENCE CHARACTERISTICS:
(A) LENGTH: 9 amino acids
(B) TYPE: amino acid
(D) TOPOLOGY: linear

(ii) MOLECULE TYPE: peptide

(xi) SEQUENCE DESCRIPTION: SEQ ID NO:5:

Asp Val Thr Ser Gly Asn Thr Val Tyr
1 5

(2) INFORMATION FOR SEQ ID NO:6:

(i) SEQUENCE CHARACTERISTICS:
(A) LENGTH: 9 amino acids
(B) TYPE: amino acid
(D) TOPOLOGY: linear

(ii) MOLECULE TYPE: peptide

(xi) SEQUENCE DESCRIPTION: SEQ ID NO:6:

Val Leu Phe Cys Ala Trp Gly Thr Leu
1 5

(2) INFORMATION FOR SEQ ID NO:7:

(i) SEQUENCE CHARACTERISTICS:
(A) LENGTH: 9 amino acids
(B) TYPE: amino acid

51

(D) TOPOLOGY: linear

(ii) MOLECULE TYPE: peptide

(xi) SEQUENCE DESCRIPTION: SEQ ID NO:7:

Cys Met His Leu Leu Leu Glu Ala Val
1 5

(2) INFORMATION FOR SEQ ID NO:8:

(i) SEQUENCE CHARACTERISTICS:
(A) LENGTH: 9 amino acids
(B) TYPE: amino acid
(D) TOPOLOGY: linear

(ii) MOLECULE TYPE: peptide

(xi) SEQUENCE DESCRIPTION: SEQ ID NO:8:

Leu Leu Leu Glu Ala Val Pro Ala Val
1 5

(2) INFORMATION FOR SEQ ID NO:9:

(i) SEQUENCE CHARACTERISTICS:
(A) LENGTH: 9 amino acids
(B) TYPE: amino acid
(D) TOPOLOGY: linear

(ii) MOLECULE TYPE: peptide

(xi) SEQUENCE DESCRIPTION: SEQ ID NO:9:

Thr Leu His Cys Asp Cys Glu Ile Leu
1 5

(2) INFORMATION FOR SEQ ID NO:10:

(i) SEQUENCE CHARACTERISTICS:
(A) LENGTH: 9 amino acids
(B) TYPE: amino acid
(D) TOPOLOGY: linear

(ii) MOLECULE TYPE: peptide

(xi) SEQUENCE DESCRIPTION: SEQ ID NO:10:

Val Leu Ser Val Asn Val Pro Asp Val
1 5

(2) INFORMATION FOR SEQ ID NO:11:

(i) SEQUENCE CHARACTERISTICS:
(A) LENGTH: 9 amino acids

(B) TYPE: amino acid
(D) TOPOLOGY: linear

(ii) MOLECULE TYPE: peptide

(xi) SEQUENCE DESCRIPTION: SEQ ID NO:11:

Asp Leu Asp Ser Thr Val Val Ala Leu
1 5

(2) INFORMATION FOR SEQ ID NO:12:

(i) SEQUENCE CHARACTERISTICS:
(A) LENGTH: 9 amino acids
(B) TYPE: amino acid
(D) TOPOLOGY: linear

(ii) MOLECULE TYPE: peptide

(xi) SEQUENCE DESCRIPTION: SEQ ID NO:12:

Trp Leu Pro Lys Ile Leu Gly Glu Val
1 5

(2) INFORMATION FOR SEQ ID NO:13:

(i) SEQUENCE CHARACTERISTICS:
(A) LENGTH: 9 amino acids
(B) TYPE: amino acid
(D) TOPOLOGY: linear

(ii) MOLECULE TYPE: peptide

(xi) SEQUENCE DESCRIPTION: SEQ ID NO:13:

Pro Leu Leu Arg Gly Leu Phe Gly Val
1 5

(2) INFORMATION FOR SEQ ID NO:14:

(i) SEQUENCE CHARACTERISTICS:
(A) LENGTH: 9 amino acids
(B) TYPE: amino acid
(D) TOPOLOGY: linear

(ii) MOLECULE TYPE: peptide

(xi) SEQUENCE DESCRIPTION: SEQ ID NO:14:

Arg Leu Gly Pro Thr Leu Met Cys Leu
1 5

(2) INFORMATION FOR SEQ ID NO:15:

- (i) SEQUENCE CHARACTERISTICS:
(A) LENGTH: 9 amino acids
(B) TYPE: amino acid
(D) TOPOLOGY: linear

(ii) MOLECULE TYPE: peptide

(xi) SEQUENCE DESCRIPTION: SEQ ID NO:15:

Leu Leu Ser Thr Gln Phe Lys Arg Leu
1 5

(2) INFORMATION FOR SEQ ID NO:16:

- (i) SEQUENCE CHARACTERISTICS:
(A) LENGTH: 9 amino acids
(B) TYPE: amino acid
(D) TOPOLOGY: linear

(ii) MOLECULE TYPE: peptide

(xi) SEQUENCE DESCRIPTION: SEQ ID NO:16:

Glu Met Gln Lys Thr Ile Thr Asp Leu
1 5

(2) INFORMATION FOR SEQ ID NO:17:

- (i) SEQUENCE CHARACTERISTICS:
(A) LENGTH: 9 amino acids
(B) TYPE: amino acid
(D) TOPOLOGY: linear

(ii) MOLECULE TYPE: peptide

(xi) SEQUENCE DESCRIPTION: SEQ ID NO:17:

Asp Leu Arg Thr Gln Ile Lys Lys Leu
1 5

(2) INFORMATION FOR SEQ ID NO:18:

- (i) SEQUENCE CHARACTERISTICS:
(A) LENGTH: 17 amino acids
(B) TYPE: amino acid
(D) TOPOLOGY: linear

(ii) MOLECULE TYPE: peptide

(xi) SEQUENCE DESCRIPTION: SEQ ID NO:18:

Met Arg Tyr Met Ile Leu Gly Leu Leu Ala Leu Ala Ala Val Cys Ser
1 5 10 15

Ala

(2) INFORMATION FOR SEQ ID NO:19:

- (i) SEQUENCE CHARACTERISTICS:
(A) LENGTH: 21 amino acids
(B) TYPE: amino acid
(D) TOPOLOGY: linear

(ii) MOLECULE TYPE: peptide

(xi) SEQUENCE DESCRIPTION: SEQ ID NO:19:

Met	Thr	Asn	Lys	Cys	Leu	Leu	Gln	Ile	Ala	Leu	Leu	Leu	Cys	Phe	Ser
1				5					10					15	
Thr	Thr	Ala	Leu	Ser											
			20												

(2) INFORMATION FOR SEQ ID NO:20:

- (i) SEQUENCE CHARACTERISTICS:
(A) LENGTH: 17 amino acids
(B) TYPE: amino acid
(D) TOPOLOGY: linear

(ii) MOLECULE TYPE: peptide

(xi) SEQUENCE DESCRIPTION: SEQ ID NO:20:

Arg	Tyr	Met	Ile	Leu	Gly	Leu	Leu	Ala	Leu	Ala	Ala	Val	Cys	Ser	Ala
1				5				10					15		
Met															

(2) INFORMATION FOR SEQ ID NO:21:

- (i) SEQUENCE CHARACTERISTICS:
(A) LENGTH: 27 amino acids
(B) TYPE: amino acid
(D) TOPOLOGY: linear

(ii) MOLECULE TYPE: peptide

(xi) SEQUENCE DESCRIPTION: SEQ ID NO:21:

Met	Arg	Tyr	Met	Ile	Leu	Gly	Leu	Leu	Ala	Leu	Ala	Ala	Val	Cys	Ser
1				5					10					15	
Ala	Arg	Pro	Arg	Pro	Glu	Gln	Glu	Pro	Leu	Pro					
			20					25							

(2) INFORMATION FOR SEQ ID NO:22:

- (i) SEQUENCE CHARACTERISTICS:
(A) LENGTH: 31 amino acids
(B) TYPE: amino acid
(D) TOPOLOGY: linear

(ii) MOLECULE TYPE: peptide

(xi) SEQUENCE DESCRIPTION: SEQ ID NO:22:

Met Thr Asn Lys Cys Leu Leu Gln Ile Ala Leu Leu Leu Cys Phe Ser
1 5 10 15
Thr Thr Ala Leu Ser Arg Pro Arg Pro Glu Gln Glu Pro Leu Pro
20 25 30

(2) INFORMATION FOR SEQ ID NO:23:

- (i) SEQUENCE CHARACTERISTICS:
(A) LENGTH: 28 amino acids
(B) TYPE: amino acid
(D) TOPOLOGY: linear

(ii) MOLECULE TYPE: peptide

(xi) SEQUENCE DESCRIPTION: SEQ ID NO:23:

Met Arg Tyr Met Ile Leu Gly Leu Leu Ala Leu Ala Ala Val Cys Ser
1 5 10 15
Ala Ala Arg Pro Arg Pro Glu Gln Glu Pro Leu Pro
20 25

(2) INFORMATION FOR SEQ ID NO:24:

- (i) SEQUENCE CHARACTERISTICS:
(A) LENGTH: 27 amino acids
(B) TYPE: amino acid
(D) TOPOLOGY: linear

(ii) MOLECULE TYPE: peptide

(xi) SEQUENCE DESCRIPTION: SEQ ID NO:24:

Arg Tyr Met Ile Leu Gly Leu Leu Ala Leu Ala Ala Val Cys Ser Ala
1 5 10 15
Met Arg Pro Arg Pro Glu Gln Glu Pro Leu Pro
20 25

(2) INFORMATION FOR SEQ ID NO:25:

- (i) SEQUENCE CHARACTERISTICS:
(A) LENGTH: 18 amino acids
(B) TYPE: amino acid
(D) TOPOLOGY: linear

56

(ii) MOLECULE TYPE: peptide

(xi) SEQUENCE DESCRIPTION: SEQ ID NO:25:

Met	Arg	Arg	Pro	Arg	Pro	Glu	Gln	Glu	Pro	Leu	Pro	Ala	Ala	Val	Cys
1				5					10					15	

Ser Ala

(2) INFORMATION FOR SEQ ID NO:26:

(i) SEQUENCE CHARACTERISTICS:

(A) LENGTH: 18 amino acids

(B) TYPE: amino acid

(D) TOPOLOGY: linear

(ii) MOLECULE TYPE: peptide

(xi) SEQUENCE DESCRIPTION: SEQ ID NO:26:

Met	Ala	Arg	Pro	Arg	Pro	Glu	Gln	Glu	Pro	Leu	Pro	Ala	Ala	Ala	Ala
1				5					10					15	

Ala Gly

We claim:

1. A substantially purified polypeptide portion of a melanoma associated antigen, MG50, comprising the amino acid sequence shown as amino acids 1187 to 1447 of
5 SEQ ID NO: 2.
2. A substantially purified T cell epitope, comprising a contiguous amino acid sequence of SEQ ID NO: 2.
3. The T cell epitope of claim 2, which is a
10 cytotoxic T cell epitope comprising 8 to 11 contiguous amino acids of SEQ ID NO: 2.
4. The T cell epitope of claim 2, comprising the amino acid sequence RPRPEQEPLP (SEQ ID NO: 4).
5. The T cell epitope of claim 2, comprising
15 the amino acid sequence DVTSGNTVY (SEQ ID NO: 5).
6. The T cell epitope of claim 2, comprising an amino acid sequence selected from the group consisting of VLFCAWGTL (SEQ ID NO: 6), CMHLLLEAV (SEQ ID NO: 7), LLLEAVPAV (SEQ ID NO: 8), TLHCDCEIL (SEQ ID NO: 9),
20 VLSVNVDPDV (SEQ ID NO: 10), DLDSTVVAL (SEQ ID NO: 11), WLPKILGEV (SEQ ID NO: 12), PLLRGLFGV (SEQ ID NO: 13), RLGPTLMCL (SEQ ID NO: 14), LLSTQFKRL (SEQ ID NO: 15), EMQKTITDL (SEQ ID NO: 16) and DLRTQIKKL (SEQ ID NO: 17).
7. The T cell epitope of claim 2, which
25 encodes a helper T cell epitope comprising 12 to 25 contiguous amino acids of SEQ ID NO: 2.
8. A substantially purified T cell epitope encoded by a nucleotide sequence contained within nucleotides 1 to 5509 of SEQ ID NO: 1.

9. A T cell epitope, comprising a cytotoxic T cell epitope having an amino acid sequence encoded by SEQ ID NO: 1, fused to a signal peptide or a functional portion thereof.

5 10. The cytotoxic T cell epitope of claim 9, wherein said signal peptide has an amino acid sequence selected from the group consisting of MRYMILGLLALAAVCSA (SEQ ID NO: 18), MTNKCLLQIALLLCFSTTALS (SEQ ID NO: 19), and RYMILGLLALAAVCSAM (SEQ ID NO: 20).

10 11. The cytotoxic T cell epitope of claim 9, comprising the amino acid sequence RPRPEQEPLP (SEQ ID NO: 4).

 12. The cytotoxic T cell epitope of claim 9, which is selected from the group consisting of:

15 MRYMILGLLALAAVCSARPRPEQEPLP (SEQ ID NO: 21),
MTNKCLLQIALLLCFSTTALS RPRPEQEPLP (SEQ ID NO: 22),
MRYMILGLLALAAVCSAARPRPEQEPLP (SEQ ID NO: 23), and
RYMILGLLALAAVCSAMRPRPEQEPLP (SEQ ID NO: 24).

 13. The cytotoxic T cell epitope of claim 9,
20 which is selected from the group consisting of:

MRRPRPEQEPLPAAVCSA (SEQ ID NO: 25), and
MARPRPEQEPLPAAAAAG (SEQ ID NO: 26).

 14. A chimeric polypeptide, comprising:

25 a) an MG50 polypeptide encoded by SEQ ID NO: 1 or an MG50 T cell epitope encoded by SEQ ID NO: 1; and

 b) a second polypeptide, which is not MG50 or an MG50 T cell epitope.

15. An antibody or an antigen binding fragment thereof that specifically binds an antigen selected from the group consisting of:

5 a) an MG50 melanoma associated antigen comprising amino acids 1187 to 1447 of SEQ ID NO: 2;

b) a peptide portion of MG50, which is encoded by a nucleotide sequence contained within nucleotides 1 to 5509 of SEQ ID NO: 1;

10 c) a cytotoxic MG50 T cell epitope encoded by a nucleotide sequence of SEQ ID NO: 1;

15 d) an MG50 T cell epitope, comprising an amino acid sequence encoded by SEQ ID NO: 1, fused to a signal peptide or a functional portion thereof; and

e) a chimeric polypeptide, comprising an MG50 polypeptide encoded by SEQ ID NO: 1 or an MG50 T cell epitope encoded by SEQ ID NO: 1.

20 16. The antibody of claim 15, wherein said peptide portion of MG50 is an MG50 T cell epitope.

17. The antibody of claim 15, which is a monoclonal antibody.

25 18. A cell expressing the antibody of claim 17.

19. An anti-idiotypic antibody, which specifically binds to the antibody of claim 17.

20. A substantially purified nucleic acid molecule, comprising nucleotides 3555 to 4336 of SEQ ID NO: 1.

21. The nucleic acid molecule of claim 20,
5 comprising nucleotides 1 to 6448 of SEQ ID NO: 1.

22. The nucleic acid molecule of claim 20,
comprising nucleotides 3555 to 6448 of SEQ ID NO: 1.

23. A substantially purified nucleic acid molecule encoding a polypeptide comprising amino acids
10 1187 to 1447 of SEQ ID NO: 2.

24. A nucleic acid molecule encoding a T cell epitope of SEQ ID NO: 2.

25. The nucleic acid molecule of claim 24,
which encodes a cytotoxic T cell epitope comprising 8
15 to 11 contiguous amino acids encoded by SEQ ID NO: 1.

26. The nucleic acid molecule of claim 25,
comprising the amino acid sequence RPRPEQEPLP (SEQ ID
NO: 4).

27. The nucleic acid molecule of claim 25,
20 comprising the amino acid sequence DVTSGNTVY (SEQ ID
NO: 5).

28. The nucleic acid molecule of claim 25,
comprising an amino acid sequence selected from the group
consisting of VLFCAGWTL (SEQ ID NO: 6), CMHLLLEAV (SEQ ID
25 NO: 7), LLLEAVPAV (SEQ ID NO: 8), TLHCDCEIL (SEQ ID NO:
9), VLSVNVDPV (SEQ ID NO: 10), DLDSTVVAL (SEQ ID NO: 11),
WLPKILGEV (SEQ ID NO: 12), PLLRGLFGV (SEQ ID NO: 13),
RLGPTLMCL (SEQ ID NO: 14), LLSTQFKRL (SEQ ID NO: 15),
EMQKTITDL (SEQ ID NO: 16) and DLRTQIKKL (SEQ ID NO: 17).

29. The nucleic acid molecule of claim 24, which encodes a helper T cell epitope comprising 12 to 25 contiguous amino acids encoded by SEQ ID NO: 1.

30. A nucleic acid molecule, comprising a
5 nucleotide sequence encoding a molecule selected from the group consisting of an MG50 T cell epitope fused to a signal peptide; and a chimeric polypeptide, comprising an MG50 polypeptide encoded by SEQ ID NO: 1 or an MG50 T cell epitope encoded by SEQ ID NO: 1.

10 31. A vector, comprising a nucleic acid molecule encoding a molecule selected from the group consisting of

a) an MG50 polypeptide, comprising amino acids 1187 to 1447 of SEQ ID NO: 2;

15 b) an MG50 T cell epitope, comprising a contiguous amino acid sequence encoded by a nucleotide sequence contained within nucleotides 1 to 5509 of SEQ ID NO: 1;

20 c) a cytotoxic MG50 T cell epitope encoded by a nucleotide sequence of SEQ ID NO: 1;

d) an MG50 T cell epitope, comprising a contiguous amino acid sequence encoded by SEQ ID NO: 1, fused to a signal peptide; and

25 e) a chimeric polypeptide, comprising an MG50 polypeptide encoded by SEQ ID NO: 1 or an MG50 T cell epitope encoded by SEQ ID NO: 1.

32. The vector of claim 31, which is an expression vector.

33. A cell containing the vector of claim 31.

34. The cell of claim 33, which is an antigen presenting cell.

35. The antigen presenting cell of claim 34,
5 which is selected from the group consisting of a dendritic cell, a mononuclear phagocytic cell, a B lymphocyte, a Langerhans cell and a human venular endothelial cell.

36. The cell of claim 33, which expresses the
10 encoded molecule on its surface.

37. A method of identifying the presence of an MG50 melanoma associated antigen in an individual, comprising the steps of:

15 a) contacting a biological sample obtained from the subject with a ligand that specifically binds MG50; and

b) detecting specific binding of the ligand and the peptide,

wherein the specific binding identifies the
20 presence of the MG50 melanoma associated antigen.

38. The method of claim 37, wherein the ligand is an antibody.

39. A method of identifying the presence in an immune response against an MG50 melanoma associated antigen in an individual, comprising the steps of:

5 a) contacting a biological sample obtained from the subject with a peptide comprising at least eight contiguous amino acids encoded by SEQ ID NO: 1; and

 b) detecting an immunoeffector function of the sample due to contact with the peptide,

10 wherein the immunoeffector function identifies the presence of an immune response against an MG50 melanoma associated antigen in the individual.

40. The method of claim 39, wherein said immunoeffector function is T cell proliferation.

15 41. The method of claim 39, wherein said immunoeffector function is specific binding by an antibody.

42. A method for producing a population of antigen presenting cells that express an MG50 T cell epitope complexed with an MHC molecule on their surfaces, comprising contacting antigen presenting cells with an MG50 melanoma associated antigen, provided said melanoma associated antigen is not CSEQPFPEHTASVQHAD (SEQ ID NO: 3).

25 43. The method of claim 42, wherein said MG50 melanoma associated antigen comprises SEQ ID NO: 2.

44. The method of claim 42, wherein said MG50 melanoma associated antigen comprises an MG50 T cell epitope encoded by SEQ ID NO: 1.

45. The method of claim 42, wherein said MG50 T cell epitope further comprises a signal peptide or a functional portion thereof.

46. The method of claim 42, wherein said
5 T cell epitope comprises the amino acid sequence RPRPEQEPLP (SEQ ID NO: 4).

47. The method of claim 42, wherein said T cell epitope comprises the amino acid sequence DVTSGNTVY (SEQ ID NO: 5).

10 48. The method of claim 42, wherein said T cell epitope comprises an amino acid sequence selected from the group consisting of VLFCAWGTL (SEQ ID NO: 6), CMHLLLEAV (SEQ ID NO: 7), LLLEAVPAV (SEQ ID NO: 8), TLHCDCEIL (SEQ ID NO: 9), VLSVNVPDV (SEQ ID NO: 10),
15 DLDSTVVAL (SEQ ID NO: 11), WLPKILGEV (SEQ ID NO: 12), PLLRGLFGV (SEQ ID NO: 13), RLGPTLMCL (SEQ ID NO: 14), LLSTQFKRL (SEQ ID NO: 15), EMQKTITDL (SEQ ID NO: 16) and DLRTQIKKL (SEQ ID NO: 17).

49. A population of antigen presenting cells
20 produced by the method of claim 40.

50. A method of producing a population of T lymphocytes specifically reactive against cancer cells expressing an MG50 melanoma associated antigen, comprising contacting T lymphocytes with the antigen
25 presenting cells of claim 49.

51. The method of claim 50, wherein the antigen presenting cells are autologous with respect to the T lymphocytes.

52. The method of claim 50, wherein the antigen presenting cells are allogeneic with respect to the T lymphocytes.

53. The method of claim 50, wherein said
5 contacting T lymphocytes with the antigen presenting cells is performed *in vitro*.

54. A population of T lymphocytes produced by the method of claim 50.

55. The method of claim 50, wherein said
10 contacting T lymphocytes with the antigen presenting cells is performed *in vivo*.

56. A method for treating an individual having a cancer containing cancer cells expressing an MG50 melanoma associated antigen, comprising administering the
15 T lymphocytes of claim 54 to the individual.

57. A method of treating an individual having a cancer containing cancer cells expressing an MG50 melanoma associated antigen, comprising administering the antigen presenting cells of claim 49 to the individual.

20 58. A method for treating an individual having a cancer containing cancer cells expressing an MG50 melanoma associated antigen, comprising administering a composition comprising an MG50 melanoma associated antigen to the individual, provided said melanoma
25 associated antigen is not CSEQPFPEHTASVTHAD (SEQ ID NO: 3).

59. The method of claim 58, wherein said composition comprises an MG50 polypeptide encoded by SEQ ID NO: 1.

60. The method of claim 58, wherein said composition comprises an MG50 T cell epitope encoded by SEQ ID NO: 1.

61. The method of claim 58, further comprising
5 administering an immunostimulatory agent to the individual.

62. The method of claim 61, wherein said immunostimulatory agent is an adjuvant.

63. The method of claim 61, wherein said
10 adjuvant is DETOX.

64. The method of claim 61, wherein said immunostimulatory agent is a cytokine.

65. The method of claim 64, wherein said cytokine is selected from the group consisting of
15 interleukin-2 and interferon- α .

66. A method for presenting an MG50 T cell epitope on the surface of an antigen presenting cell, comprising contacting the antigen presenting cell with a nucleic acid molecule encoding a molecule selected from
5 the group consisting of:

a) an MG50 polypeptide, comprising amino acids 1187 to 1447 of SEQ ID NO: 2;

b) an MG50 T cell epitope, comprising a contiguous amino acid sequence of SEQ ID NO: 2;

10 c) a cytotoxic MG50 T cell epitope encoded by SEQ ID NO: 1;

d) an MG50 T cell epitope, comprising a contiguous amino acid sequence encoded by SEQ ID NO: 1, fused to a signal peptide.

15 67. A population of antigen presenting cells produced by the method of claim 66.

68. A method for treating an individual having cancer cells expressing an MG50 melanoma associated antigen, comprising administering the antigen presenting
20 cells of claim 65 to the individual.

69. A method for treating an individual having cancer cells expressing an MG50 melanoma associated antigen, comprising administering to the individual a composition comprising a nucleic acid molecule encoding a molecule selected from the group consisting of:

a) an MG50 polypeptide, comprising amino acids 1187 to 1447 SEQ ID NO: 2;

b) an MG50 T cell epitope, comprising a contiguous amino acid sequence of SEQ ID NO: 2;

10 c) a cytotoxic MG50 T cell epitope encoded by SEQ ID NO: 1; and

d) an MG50 T cell epitope, comprising a contiguous amino acid sequence encoded by SEQ ID NO: 1, fused to a signal peptide.

1 AGCCGGCCGT GGTGGCTCCG TCGTCCGAG CGTCCGTCCG CGCCGTCGGC CATGGCCAAG
61 CGCTCCAGGG GCCCGGGCG CCGCTGCCTG TTGGCGCTCG TGCTGTTCTG CGCCTGGGGG
121 ACGCTGGCCG TGGTGGCCCA GAAGCCGGGC GCAGGGTGTC CGAGCCGCTG CCTGTGCTTC
181 CGCACCACCG TGCGCTGCAT GCATCTGCTG CTGGAGGCCG TGCCCGCCGT GGCGCCGAG
241 ACCTCCATCC TAGATCTTCG CTTTAACAGA ATCAGAGAGA TCCAACCTGG GGCATTACAG
301 CGGCTGAGGA ACTTGAACAC ATTGCTTCTC AATAATAATC AGATCAAGAG GATACCTAGT
361 GGAGCATTTG AAGACTTGGA AAATTTAAAA TATCTCTATC TGTACAAGAA TGAGATCCAG
421 TCAATTGACA-GGCAAGCATT TAAGGGACTT GCCTCTCTAG AGCAACTATA CCTGCACTTT
481 AATCAGATAG AAACCTTGGA CCCAGATTCG TTCCAGCATC TCCCGAAGCT CGAGAGGCTA
541 TTTTGCATA ACAACCGGAT TACACATTTA GTTCCAGGGA CATTTAATCA CTTGGAATCT
601 ATGAAGAGAT TGCGACTGGA CTCAAACACA CTTCACTGCG ACTGTGAAAT CCTGTGGTTG
661 GCGGATTTGC TGAAAACCTA CGCGGAGTCG GGAACGCGC AGGCAGCGGC CATCTGTGAA
721 TATCCCAGAC GCATCCAGGG ACGCTCAGTG GCAACCATCA CCCCGBAAGA GCTGAACTGT
781 GAAAGGCCCC GGATCACCTC CGAGCCCCAG GACGCAGATG TGACCTCGGG GAACACCGTG
841 TACTTCACCT GCAGAGCCGA AGGCAACCCC AAGCCTGAGA TCATCTGGCT GCGAAACAAT
901 AATGAGCTGA GCATGAAGAC AGATTCCCGC CTAAACTTGC TGGACGATGG GACCCTGATG
961 ATCCAGAACA CACAGGAGAC AGACCAGGGT ATCTACCAGT GCATGGCAA GAACGTGGCC
1021 GGAGAGGTGA AGACGCAAGA GGTGACCCTC AGGTACTTCG GGTCTCCAGC TCGACCCACT
1081 TTTGTAATCC AGCCACAGAA TACAGAGGTG CTGGTTGGGG AGAGCGTCAC GCTGGAGTGC
1141 AGCGCCACAG GCCACCCCCC GCCGCGGATC TCCTGGACGA GAGGTGACCG CACACCCTTG
1201 CCAGTTGACC CGCGGGTGAA CATCACGCCT TCTGGCGGGC TTTACATACA GAACGTGCTA
1261 CAGGGGGACA GCGGAGAGTA TGCGTGCTCT GCGACCAACA ACATTGACAG CGTCCATGCC
1321 ACCGCTTTCA TCATCGTCCA GGCTCTTCCT CAGTTCACTG TGACGCCTCA GGACAGAGTC
1381 GTTATTGAGG GCCAGACCGT GGATTTCAG TGTGAAGCCA AGGGCAACCC GCCGCCGTC
1441 ATCGCCTGGA CCAAGGGAGG GAGCCAGCTC TCCGTGGACC GGCGGCACCT GGTCTGTCA
1501 TCGGGAACAC TTAGAATCTC TGGTGTTGCC CTCCACGACC AGGGCCAGTA CGAATGCCAG

FIG. 1A

2 / 6

1561 GCTGTCAACA TCATCGGCTC CCAGAAGGTC GTGGCCCACC TGACTGTGCA GCCCAGAGTC
1621 ACCCCAGTGT TTGCCAGCAT TCCCAGCGAC ACAACAGTGG AGGTGGGCGC CAATGTGCAG
1681 CTCCCGTGCA GCTCCAGGG CGAGCCCGAG CCAGCCATCA CCTGGAACAA GGATGGGGTT
1741 CAGGTGACAG AAAGTGGAAT ATTTACATC AGCCCTGAAG GATTCTTGAC CATCAATGAC
1801 GTTGGCCCTG CAGACGCAGG TCGCTATGAG TGTGTGGCCC GGAACACCAT TGGGTCCGCC
1861 TCGGTGAGCA TGGTGCTCAG TGTGAACGTT CCTGACGTCA GTCGAAATGG AGATCCGTTT
1921 GTAGCTACCT CCATCGTGGA AGCGATTGCG ACTGTTGACA GAGCTATAAA CTCAACCCGA
1981 ACACATTTGT TTGACAGCCG TCCTCGTTCT CCAAATGATT TGCTGGCCTT GTTCCGGTAT
2041 CCGAGGGATC CTTACACAGT TGAACAGGCA CGGGCGGGAG AAATCTTTGA ACGGACATTG
2101 CAGCTCATTC AGGAGCATGT ACAGCATGGC TTGATGGTCG ACCTCAACGG AACAAGTTAC
2161 CACTACAACG ACCTGGTGTC TCCACAGTAC CTGAACCTCA TCGCAAACCT GTCGGGCTGT
2221 ACCGCCCACC GGCGCGTGAA CAACTGCTCG GACATGTGCT TCCACCAGAA GTACCGGACG
2281 CACGACGGCA CCTGTAACAA CCTGCAGCAC CCCATGTGGG GCGCCTCGCT GACCGCCTTC
2341 GAGCGCCTGC TGAAATCCGT GTACGAGAAT GGCTTCAACA CCCCTCGGGG CATCAACCCC
2401 CACCGACTGT ACAACGGGCA CGCCCTTCCC ATGCCGCGCC TGGTGTCCAC CACCCTGATC
2461 GGGACGGAGA CCGTCACACC CGACGAGCAG TTCACCCACA TGCTGATGCA GTGGGGCCAG
2521 TTCCTGGACC ACGACCTCGA CTCCACGGTG GTGGCCCTGA GCCAGGCACG CTTCTCCGAC
2581 GGACAGCACT GCAGCAACGT GTGCAGCAAC GACCCCCCTT GCTTCTCTGT CATGATCCCC
2641 CCCAATGACT CCCGGGCCAG GAGCGGGGCC CGCTGCATGT TCTTCGTGCG CTCCAGCCCT
2701 GTGTGCGGCA GCGGCATGAC TTCGCTGCTC ATGAACTCCG TGTACCCGCG GGAGCAGATC
2761 AACCAGCTCA CTCCTACAT CGACGCATCC AACGTGTACG GGAGCACGGA GCATGAGGCC
2821 CGCAGCATCC GCGACCTGGC CAGCCACCGC GGCCTGCTGC GGCAGGGCAT CGTGCAGCGG
2881 TCCGGGAAGC CGCTGCTCCC CTTGCGCCACC GGGCCGCCCA CGGAGTGCAT GCGGGACGAG
2941 AACGAGAGCC CCATCCCCTG CTTCTGGCC GGGGACCACC GCGCCAACGA GCAGCTGGGC
3001 CTGACCAGCA TGCACACGCT GTGGTTCCGC GAGCACAACC GCATTGCCAC GGAGCTGCTC
3061 AAGCTGAACC CGCACTGGGA CGGCGACACC ATCTACTATG AGACCAGGAA GATCGTGGGT
3121 GCGGAGATCC AGCACATCAC CTACCAGCAC TGGCTCCCGA AGATCCTGGG GGAGGTGGGC

FIG. 1B

3181 ATGAGGACGC TGGGAGAGTA CCACGGCTAC GACCCCGGCA TCAATGCTGG CATCTTCAAC
3241 GCCTTCGCCA CCGCGGCCTT CAGGTTTGGC CACACGCTTG TCAACCCACT GCTTTACCGG
3301 CTGGACGAGA ACTTCCAGCC CATTGCACAA GATCACCTCC CCCTTCACAA AGCTTTCTTC
3361 TCTCCCTTCC GGATTGTGAA TGAGGGCGGC ATCGATCCGC TTCTCAGGGG GCTGTTCCGG
3421 GTGGCGGGGA AAATGCGTGT GCCCTCGCAG CTGCTGAACA CGGAGCTCAC GGAGCGGCTG
3481 TTCTCCATGG CACACACGGT GGCTCTGGAC CTGGCGGCCA TCAACATCCA GCGGGGCCGG
3541 GACCACGGGA TCCCACCCTA CCACGACTAC AGGGTCTACT GCAATCTATC GGCGGCACAC
3601 ACGTTCGAGG ACCTGAAAAA TGAGATTAAA AACCTGAGA TCCGGGAGAA ACTGAAAAGG
3661 TTGTATGGCT CGACACTCAA CATCGACCTG TTTCCGGCGC TCGTGGTGGG GGACCTGGTG
3721 CCTGGCAGCC GGCTGGGCCC CACCCTGATG TGTCTTCTCA GCACACAGTT CAAGCGCCTG
3781 CGAGATGGGG ACAGGTTGTG GTATGAGAAC CCTGGGGTGT TCTCCCCGGC CCAGCTGACT
3841 CAGATCAAGC AGACGTCGCT GGCCAGGATC CTATGCGACA ACGCGGACAA CATCACCCGG
3901 GTGCAGAGCG ACGTGTTTCT GGTGGCGGAG TTCCCTCAGC GCTACGGCAG CTGTGACGAG
3961 ATCCCCAGGG TGGACCTCCG GGTGTGGCAG GACTGCTGTG AAGACTGTAG GACCAGGGGG
4021 CAGTTCAATG CCTTTTCCTA TCATTTCCTG GGCAGACGGT CTCTTGAGTT CAGCTACCAG
4081 GAGGACAAGC CGACCAAGAA AACAAGACCA CGGAAAATAC CCAGTGTTGG GAGACAGGGG
4141 GAACATCTCA GCAACAGCAC CTCAGCCTTC AGCACACGCT CAGATGCATC TGGGACAAAT
4201 GACTTCAGAG AGTTTGTCTT GGAAATGCAG AAGACCATCA CAGACCTCAG AACACAGATA
4261 AAGAACTTG AATCACGGCT CAGTACCACA GAGTGCGTGG ATGCCGGGGG CGAATCTCAC
4321 GCCAACAACA CCAAGTGGAA AAAAGATGCA TGCACCATTG GTGAATGCAA AGACGGGCAG
4381 GTCACCTGCT TCGTGGAAGC TTGCCCCCCT GCCACCTGTG CTGTCCCCGT GAACATCCCA
4441 GGGGCCTGCT GTCCAGTCTG CTTACAGAAG AGGGCGGAGG AAAAGCCCTA GGCTCCTGGG
4501 AGGCTCCTCA GAGTTTGTCT GCTGTGCCAT CGTGAGATCG GGTGGCCGAT GGCAGGGAGC
4561 TCGGGAATCG AGACCAGGAA ACACCCAGAA CTCGTGACAT TTCATGACAA CGTCCAGCTG
4621 GTGCTGTTAC AGAAGGCAGT GCAGGAGGCT TCCAACCAGA GCATCTGCGG AGAAGGAGGC
4681 ACAGCAGGTG CCTGAAGGGA AGCAGGCAGG AGTCCTAGCT TCACGTTAGA CTTCTCAGGT
4741 TTTTATTTAA TTCTTTTAAA ATGAAAAATT GGTGCTACTA TTAAATTGCA CAGTTGAATC

FIG. 1C

4801 ATTTAGGCGC CTAAATTGGT TTTGCCCTCCC AACACCATT TTTTTTAAAT AAAGCAGGAT
4861 ACCTCTATAT GTCAGCCTTG CCTTGTTTCTAG ATGCCAGGAG CCGGCAGACC TGTCACCCGC
4921 AGGTGGGGTG AGTCTCGGAG CTGCCAGAGG GGCTCACC GAATCGGGGTT CCATCACAAG
4981 CTATGTTTAA AAAGAAAATT GGTGTTTGGC AAACGGAACA GAACCTTTGA TGAGAGCGTT
5041 CACAGGGACA CTGTCTGGGG GTGCAGTGCA AGCCCCGGC CTCTTCCCTG GGAACCTCTG
5101 AACTCCTCCT TCCTCTGGGC TCTCTGTAAC ATTTACCAC ACGTCAGCAT CTAATCCCAA
5161 GACAAACATT CCCGCTGCTC GAAGCAGCTG TATAGCCTGT GACTCTCCGT GTGTCAGCTC
5221 CTTCCACACC TGATTAGAAC ATTCATAAGC CACATTTAGA AACAGATTTG CTTTCAGCTG
5281 TCACTTGCAC ACATACTGCC TAGTTGTGAA CCAAATGTGA AAAAACCTCC TTCATCCCAT
5341 TGTGTATCTG ATACCTGCCG AGGGCCAAGG GTGTGTGTTG ACAACGCCGC TCCAGCCGG
5401 CCCTGGTTGC GTCCACGTCC TGAACAAGAG CCGCTTCCGG ATGGCTCTTC CCAAGGGAGG
5461 AGGAGCTCAA GTGTCGGGAA CTGTCTAACT TCAGGTTGTG TGAGTGCGTT AAAAAAAAAA
5521 AAAAAAAAAA GAATCCCTAT ACCTCATTTG TATTTTAAAT ATGCGTGATG TTTTATGAAA
5581 TTGTGTCCAT TTTTATAGGTA TTAGATATGG CAGAAAAACC ATTTCCACTA TGCAAAGTTC
5641 TTTTAGACGT CAGTGAAAAT CAACTCTCAT ACCTCATGGG TCTCTCTTTA ATTGACCAAA
5701 ACCTTCCATT TTTCTCTTAA ATACAAAGCG ATCTGTGTTC TGAGCAACCT TTCCCCGAAC
5761 ACACAGCTTC AGTGCAGCAC GCTGACCTGA GTATCCACCA GGTGCCAGGC ACAGTTGCTG
5821 GGCNNACGGA GGCACCAAGG TCCGGGCCAC CTGCCCCGAG GCAAGGCCCA GCTGAGGTGG
5881 TGGGAGGGGA GCCCCTGAGG TCAGGGGCCG TTTCGGTTCA GGGTGGCAGG TGTCCAGCAC
5941 TGGGGTATGG CGTCGAGGCT TCCATGGCGT GGGGGAGGCC AGCTTCCTTC TGACAGGATG
6001 GGCGCATACA GTGCCTGGTG TGATTTGTGC ACAACCCGTG TTCCAGGTGC ACATCCTCCC
6061 AAGGAGACAC CCAGACCCTT CCAGCACGGG CCGGCCAAGT TGCTGCGGCG GAGGCAGCAT
6121 TTCAGCTGTG AGGAAGGTCA TTGGATTCTG GTGTTTTATC TGTAATAATG GTTGTCTTAA
6181 CTTCTTAACT CATATTGGTA AGTGATTGAT AAAAATTGGT TGGTGTCTTC ATGACATGTG
6241 GACTTCTNTT GNATAGAAGT CAAATGTAGT GACAATTTGT GGAAGAGATT CTTGTCAAAG
6301 TGAAATAGGA AATGTGTAAG TTCGTCTAAA AGCTGATGGT TATGTAAGTT GCTCAGGCAC
6361 TCAGATGACA GCAGATTCTG GGTCTCTGGA GTGTTCTGTG CCTCTTACAT GCCCTGGAGG

FIG. 1D

6421 CCTCATGGTC TCAGTGCTGA GCGGGCACAC CTGTAGCACA CCTGCGTAAT GTGCGGTCTG
6481 GGCCAGTCAC AAGGAATTGT GTTGTCTAAN CCAAAGGGGG AAGCTDACTG TGTATTACCA
6541 AAAAAAATTC TGTAATNCAA ACCNAAATGT CTGCGGAATC ACCAGTTTGA TACTCTCTGT
6601 AATCAGAGCA GTNGNCTGAG GCGGGNCAGT NCCTGGGTGA ACGTGTCTAG CAGCCACTGT
6661 GGGGGATCGC TGTAACAGGA GTGGAATGTA CATATTTATT TACTTTTCTA ACTGCTCCAA
6721 CAGCCAAATG CCTTTTTTAT GACCATTGTA TTCAGTTCAT TACCAAAGAA ATGTTTGCAC
6781 TTTGTAATGA TGCCTTTCAG TTCAAATAAA TGGGTCACAT TTTCAAATGG AAAAAAAAAA
6841 AAAAAAA

FIG. 1E

1 SRPWWLRASE RPSAPSAMAK RSRGPGRRL LALVLFCAWG TLAVVAQKPG AGCPSRCLCF 61
 RTTVRCMHLL LEAVPAVAPQ TSILDLREN IREIQPGA FR RLRNLNTLLL NNNQIKRIPS 121
 GAFEDLENLK YLYLYKNEIQ SIDRQAFKGL ASLEQLYLHF NQIETLDPDS FQHLPKLERL 181
 FLHNNRITHL VPGTFNHLES MKRLRLDSNT LHCDCEILWL ADLLKTYAES GNAQAAAICE 241
 YPRRIQGRSV ATITPEELNC ERPRITSEPQ DADVTSGNTV YFTCRAEGNP KPEIIWLRNN 301
 NELSMKTD SR LNLDDGTLM IQNTQETDQG IYQCMKNVA GEVKTQEVTL RYFGSPARPT 361
 FVIQPQNT EV LVGESVTLEC SATGHPPPRI SWTRGDRTPL PVDPRVNITP SGGLYIQNVV 421
 QGDSGEYACS ATNNIDSVHA TAFIIVQALP QFTVTPQDRV VIEGQTVDFQ CEAKGNPPP V 481
 IAWTKGGSQ L SVDRRHVLVS SGTLRISGVA LHDQGQYECQ AVNIIGSQKV VAHLTVQPRV 541
 TPVFASIPSD TTVEVGANVQ LPCSSQGEPE PAITWNKDG V QVTESGKFHI SPEGFLTIND 601
 VGPADAGRYE CVARNTIGSA SVSMVLSVNV PDVSRNGDPF VATSIVEAIA TVDRAINSTR 661
 THLFDSRPRS PNDLLALFRY PRDPYTVEQA RAGEIFERTL QLIQEHVQHG LMVDLNGTSY 721
 HYNDLVSPQY LNLIANLSGC TAHRRVNNCS DMCFHQKYRT HDGTCNNLQH PMWGASLTAF 781
 ERLKSVYEN GFNTPRGINP HRLYNHALP MPRLVSTTLI GTETVTPDEQ FTHMLMQWGQ 841
 FLDHDL DSTV VALSQARFSD GQHCSNVCSN DPPCF SVMIP PNDSRARS GA RCMFFVRSSP 901
 VCGSGMTSLL MNSVYPREQI NQLTSYIDAS NVYGSTEHEA RSIRDLASHR GLLRQGIVQR 961
 SGKPLLPFAT GPPTECMRDE NESPIPCFLA GDHRANEQLG LTSMHTLWFR EHNRIATELL 1021
 KLNPHWDGDT IYYETRKIVG AEIQHITYQH WLPKILGEVG MRTLG EYHGY DPGINAGIFN 1081
 AFATAAFRFG HTLVNPLLYR LDENFOPIAQ DHLPLHKAFF SPFRIVNEG IDPLLRGLFG 1141
 VAGKMRVPSQ LLNTE LTERL FSMAHTVALD LAAINIQRGR DHGIPPYHDY RVCNLSAAH 1201
 TFEDLKNEIK NPEIREKLKR LYGSTLNIDL FPALVVEDLV PGSRLGPTLM CLLSTQFKRL 1261
 RDGDRLWYEN PGVFSPAQLT QIKQTS LARI LCDNADNITR VQSDVFRVAE FPHGYGSCDE 1321
 IPRVDLRVWQ DCCEDCRTRG QFNAFSYHFR GRRSLEFSYQ EDKPTKKTRP RKIPSVGRQG 1381
 EHLSNSTSAF STRSDASGTN DFREFVLEMQ KTITDLRTQI KKLESRLSTT ECVDAGGES H 1441
 ANNTKWKKDA CTICECKDGQ VTCFVEACPP ATCAVPVNIP GACCPVCLQK RAE EKP

FIG. 2

INTERNATIONAL SEARCH REPORT

International application No.
PCT/US98/11533

A. CLASSIFICATION OF SUBJECT MATTER

IPC(6) :A61K 38/00, 45/05; C07K 14/00, 14/82; C12N 15/00; C12Q 1/00

US CL :Please See Extra Sheet.

According to International Patent Classification (IPC) or to both national classification and IPC

B. FIELDS SEARCHED

Minimum documentation searched (classification system followed by classification symbols)

U.S. : 424/185.1, 277; 530/300, 328, 350, 395; 435/7.1, 69.3, 172.1

Documentation searched other than minimum documentation to the extent that such documents are included in the fields searched

Electronic data base consulted during the international search (name of data base and, where practicable, search terms used)

EMBL, GENBANK nucleic acid sequence search of SEQ ID NO: 1; GENESEQ32 amino acid sequence search of SEQ ID NO: 2 and residues 1187-1447 of SEQ ID NO: 2 and SEQ ID NOS: 3-26.

C. DOCUMENTS CONSIDERED TO BE RELEVANT

Category*	Citation of document, with indication, where appropriate, of the relevant passages	Relevant to claim No.
Y	WO 96/40907 A1 (GENETICS INSTITUTE, INC.) 19 December 1996, see entire document.	1-14, 20-32, 39-41, 50-56
Y	WO 94/21680 A1 (THE GOVERNMENT OF THE UNITED STATES OF AMERICA) 29 September 1994, see entire document.	1-14, 20-32, 39-41, 50-56

☐ Further documents are listed in the continuation of Box C. ☐ See patent family annex.

* Special categories of cited documents:	*T* later document published after the international filing date or priority date and not in conflict with the application but cited to understand the principle or theory underlying the invention
A document defining the general state of the art which is not considered to be of particular relevance	*X* document of particular relevance; the claimed invention cannot be considered novel or cannot be considered to involve an inventive step when the document is taken alone
E earlier document published on or after the international filing date	*Y* document of particular relevance; the claimed invention cannot be considered to involve an inventive step when the document is combined with one or more other such documents, such combination being obvious to a person skilled in the art
L document which may throw doubt on priority claim(s) or which is cited to establish the publication date of another citation or other special reason (as specified)	* & * document member of the same patent family
O document referring to an oral disclosure, use, exhibition or other means	
P document published prior to the international filing date but later than the priority date claimed	

Date of the actual completion of the international search

31 AUGUST 1998

Date of mailing of the international search report

13 OCT 1998

Name and mailing address of the ISA/US
Commissioner of Patents and Trademarks
Box PCT
Washington, D.C. 20231

Facsimile No. (703) 305-3230

Authorized officer

Thomas Cunningham

Telephone No. (703) 308-0196

INTERNATIONAL SEARCH REPORT

International application No.
PCT/US98/11533

Box I Observations where certain claims were found unsearchable (Continuation of item 1 of first sheet)

This international report has not been established in respect of certain claims under Article 17(2)(a) for the following reasons:

1. ☐ Claims Nos.:
because they relate to subject matter not required to be searched by this Authority, namely:
2. ☐ Claims Nos.:
because they relate to parts of the international application that do not comply with the prescribed requirements to such an extent that no meaningful international search can be carried out, specifically:
3. ☐ Claims Nos.:
because they are dependent claims and are not drafted in accordance with the second and third sentences of Rule 6.4(a).

Box II Observations where unity of invention is lacking (Continuation of item 2 of first sheet)

This International Searching Authority found multiple inventions in this international application, as follows:

Please See Extra Sheet.

1. ☐ As all required additional search fees were timely paid by the applicant, this international search report covers all searchable claims.
2. ☐ As all searchable claims could be searched without effort justifying an additional fee, this Authority did not invite payment of any additional fee.
3. ☒ As only some of the required additional search fees were timely paid by the applicant, this international search report covers only those claims for which fees were paid, specifically claims Nos.:
1-14, 20-32, 39-41, 50-56
4. ☐ No required additional search fees were timely paid by the applicant. Consequently, this international search report is restricted to the invention first mentioned in the claims; it is covered by claims Nos.:

Remark on Protest

- ☐ The additional search fees were accompanied by the applicant's protest.
☐ No protest accompanied the payment of additional search fees.

INTERNATIONAL SEARCH REPORT

International application No.
PCT/US98/11533

A. CLASSIFICATION OF SUBJECT MATTER:

US CL :

424/93.71, 185.1, 277.1; 530/300, 328, 350, 395; 435/7.1, 69.3, 172.1; 536/23.5

BOX II. OBSERVATIONS WHERE UNITY OF INVENTION WAS LACKING

This ISA found multiple inventions as follows:

This application contains the following inventions or groups of inventions which are not so linked as to form a single inventive concept under PCT Rule 13.1. In order for all inventions to be searched, the appropriate additional search fees must be paid.

Group I, claim(s) 1-14 and 39-41, drawn to peptides/polypeptides and the first method of using said peptides/polypeptides.

Group II, claim(s) 15-19, drawn to antibodies and hybridoma cell lines.

Group III, claim(s) 20-32, drawn to nucleic acid and nucleic acid vectors.

Group IV, claims 33-36 and 49, drawn to antigen presenting cells.

Group V, claims 37-38, drawn to methods of detecting using ligands such as antibodies.

Group VI, claims 42-48, drawn to a method for producing antigen presenting cells.

Group VII, claims 50-55, drawn to methods of making T cells.

Group VIII, claims 56, drawn to method of treatment using T cells.

Group IX, claims 57, drawn to method of treatment using antigen presenting cells.

Group X, claims 58-65, drawn to method of treatment using MG50 peptides/polypeptides.

Group I of this application contains claims directed to more than one species of the generic invention. These species are deemed to lack Unity of Invention because they are not so linked as to form a single inventive concept under PCT Rule 13.1 and lack a common core structure. In order for more than one species within Group I to be searched, the appropriate additional search fees must be paid. The species are as follows:

Species I(a) (residues 1187-1447 SEQ ID NO:2), species I(b) (SEQ ID NO:2), species I(c) (SEQ ID NO:3), species I(d) (SEQ ID NO:4), species I(e) (SEQ ID NO:5), species I(f) (SEQ ID NO:6), species I(g) (SEQ ID NO:7), species I(h) through species I(z) corresponding to SEQ ID NOS 8-26, respectively. The peptide which comprises amino acid residues 1187-1447 of SEQ ID NO:2 (species I(a)) is the first mentioned species and will be searched with no additional charge. Claims 1-14 and 39-41 all appear to encompass each species due to the open claim language.

Group II of this application contains claims directed to more than one species of the generic invention. These species are deemed to lack Unity of Invention because they are not so linked as to form a single inventive concept under PCT Rule 13.1 and lack a common core structure. In order for more than one species within Group I to be searched, the appropriate additional search fees must be paid. The species are as follows antibodies which bind to the following species of peptides:

Species II(a) (antibodies binding to residues 1187-1447 SEQ ID NO:2), species I(b) (antibodies binding to SEQ ID NO:2), species II(c) (antibodies binding to SEQ ID NO:3), species II(d) (antibodies binding to SEQ ID NO:4), species II(e) (antibodies binding to SEQ ID NO:5), species II(f) (antibodies binding to SEQ ID NO:6), species II(g) (antibodies binding to SEQ ID NO:7), species II(h) through species II(z) corresponding to antibodies binding to SEQ ID NOS 8-26, respectively. Species II(a) peptide which comprises amino acid residues 1187-1447 of SEQ ID NO:2 (species I(a)) is the first mentioned species for Group II and will be searched with no additional charge upon payment for Group II. Claims 15-19 all appear to encompass each species due to open claim language.

Group III of this application contains claims directed to more than one species of the generic invention. These species are deemed to lack Unity of Invention because they are not so linked as to form a single inventive concept under PCT Rule 13.1 and lack a common core structure. In order for more than one species within Group I to be searched, the appropriate additional search fees must be paid. The species are as follows nucleic acids encoding the following species of peptides:

Species III(a) (nucleic acids comprising nucleotides 3555-4336, 1-6448, or 3555-6448 of SEQ ID NO:1; species III(b) nucleic acids encoding 1187-1447 SEQ ID NO:2), species III(c) nucleic acids encoding SEQ ID NO:2, species III(d) nucleic acids encoding SEQ ID NO:3, species III(e) nucleic acids encoding SEQ ID NO:4, species III(f) nucleic acids encoding SEQ ID NO:5, species III(g) nucleic acids encoding SEQ ID NO:6, species III(h) nucleic acids encoding SEQ ID NO:7, species III(i) through species III(z) and III(aa) corresponding to nucleic acids encoding SEQ ID NOS 8-26, respectively. Species III(a) is the first mentioned species for Group III and will be searched with no additional charge.

INTERNATIONAL SEARCH REPORT

International application No.
PCT/US98/11533

upon payment for Group III. Claims 20-32 all appear to encompass each species due to the open claim language.

The claims are deemed to correspond to the species listed above in the following manner:

The following claims are generic: for Group I: 1-14 and 39-41. For Group II: 15-19; for Group III: 20-32.

The inventions listed as Groups I-X do not relate to a single inventive concept under PCT Rule 13.1 because, under PCT Rule 13.2, they lack the same or corresponding special technical features for the following reasons:

Each group is directed to products which lack the same or corresponding technical features because they have different structures and functions, e.g. polypeptides, antibodies, nucleic acids, T cells, or antigen presenting cells. The groups directed to methods lack corresponding technical features because they recited different method steps are require the use of products which lack corresponding technical features.

The species listed above do not relate to a single inventive concept under PCT Rule 13.1 because, under PCT Rule 13.2, the species lack the same or corresponding special technical features for the following reasons:

Each species of peptide/polypeptide lacks the same or corresponding technical features because they have different structures and functions, e.g. correspond to different epitopes.